New Computationally Efficient Quantum Chemical Models that Capture Static and Dynamic Correlation Separately

Chris Stinson

A thesis presented for the degree of

Doctor of Philosophy in Chemistry

Department of Chemistry

University of Canterbury

October, 2020

Contents

Abstract

This thesis presents a new family of computationally-efficient quantum chemical methods designed to capture static and dynamic correlation energies separately without double counting. Statically correlated systems require more than one Slater determinant to qualitatively describe the electronic wavefunction. A major challenge for modelling statically correlated systems is finding molecular orbitals that are simultaneously suitable to describe all relevant electronic configurations. Conventional multi-reference SCF methods overcome this problem by simultaneously optimizing molecular orbital and configuration interaction coefficients, but this procedure is computationally intensive and selecting the relevant configurations and orbitals is far from straightforward. Spin-flip non-orthogonal configuration interaction methods allow these processes to be decoupled by generating semi-optimized orbitals for excited-state configurations using high-spin reference Hartree-Fock calculations. The key insight of this thesis is that these expansions can be severely truncated in a physically-motivated manner to provide minimal determinant models that capture only static correlation, and to which a very simple second-order perturbation theory correction can be applied to recover the remaining dynamic correlation energy. Our minimal-determinant SF-NOCI and SF-NOCI-PT2 methods are applied to two simple model problems in which static correlation effects are important - dissociating LiH and twisting ethylene. SF-NOCI gives qualitatively-correct wavefunctions while SF-NOCI-PT2 energies show close agreement with experiment for ground-state dissociation energies and torsional barrier heights, and limited agreement for excited-state transition energies. This demonstrates the importance of optimizing orbitals for excited determinants. Even approximately optimized orbitals can substantially improve model accuracy while reducing complexity and computational cost. Accounting for dynamic and static correlation effects separately allows efficient models to be developed or used for both.

Chapter 1

Introduction

1.1 Molecular Orbital Theory

1.1.1 Qualitative Molecular Orbitals

The behaviour of electrons and protons is one of the core concerns of chemistry and many of the key ideas in the discipline amount to different ways to think about how electrons are distributed within and between molecules. The behaviour of electrons is described by the electronic Schrödinger equation and its solution: the electronic wavefunction. However the electronic wavefunction is a complex mathematical object which involves the position and spin coordinates of all the electrons in the system and, like any n-body problem, the Schrödinger equation is insolvable for all but the simplest chemical systems.[1] Consequently, chemists have devised strategies for thinking about electron behaviour that avoid the full complexity of the wavefunction by providing useful heuristics for understanding chemical behaviour. These include the concept of electronegativity[2] and the "curly arrow" methods[3] common in organic chemistry.

Of these strategies, one of the most fundamental to understanding chemistry is the orbital model, which takes the multi-electron wavefunction and factorizes it into much simpler single-electron functions in three dimensional space that are much

easier to understand and visualize.[4] This model derives from the solutions to the Schrödinger equation for the hydrogen atom, a simple one-electron system that gives relatively simple, easy to understand solutions, which also form a reasonable approximation for understanding the behaviour of electrons in larger atoms. These hydrogen-like atomic orbitals (AOs) are then commonly combined to construct molecular orbitals (MOs), which can be used to understand the behaviour of electrons in molecules, including bonding, photochemistry *etc.*[5] This introduces one of the fundamental strategies in all quantum chemical models - using simpler pictures of the wavefunction as a basis to build up representations with higher accuracy or explanatory power.

In molecular orbital theory, the MOs are constructed as a linear combination of AOs (LCAO), *i.e.* a sum of differently weighted AOs. Commonly the resulting MOs are justified on the basis of symmetry and energetic arguments *i.e.* AOs will only mix with other orbitals of the same symmetry and will mix more strongly with AOs of similar energy.[6] However in cases where multiple non-equivalent AOs contribute to the same MO (*e.g.* s and p_z AOs both contribute to the σ MOs of a diatomic molecule) this approach can only say that both AOs will contribute, not quantify their relative contributions. Consequently, symmetry-based MO descriptions are typically only useful in *post-hoc* justification of molecular properties.

Nonetheless, the orbital picture of electron behaviour has been incredibly effective - to the extent that it forms the basis of how most chemists think about electronic structure. At a qualitative level it provides an intuitive explanation of how and why chemical bonds form[7] and is an invaluable tool in understanding and interpreting spectra, and predicting molecular reactivity.[8,9]

1.1.2 Hartree-Fock Theory

Hartree-Fock (HF) theory provides the quantitative basis for all qualitative MO theories, *e.g.* Ligand Field theory, Valence Shell Electron-Pair Repulsion theory

(VSEPR), and Walsh diagrams.[10] It constructs the MOs of qualitative molecular orbital theory as a linear combination of atomic orbital basis functions drawn from an atomic orbital basis set. These orbitals are then used to construct a Slater determinant,[11] which represents the total HF wavefunction as an appropriately antisymmetrized product of the occupied molecular orbitals. This ensures that the Pauli exclusion principle is satisfied, *i.e.* that each MO can only be occupied by a pair of opposite-spin electrons, and that the wavefunction changes sign if electrons are moved between MOs.[4] The occupied MOs are chosen according to the Aufbau principle, "filling up" the molecular orbitals in order of ascending energy. The MO coefficients are optimized by minimizing the electronic energy of the molecule as a function of the contributions of the AOs to each occupied MO. This process yields values for the energy and composition of each MO as well as the total electronic energy of the molecule.

The HF method is variational, which means that it gives an upper bound on the true energy of the system, and systematic improvements to the modelling of the wavefunction or the flexibility of the orbitals will cause the energy to decrease towards the true value. The difference between the HF energy and the actual energy is referred to as the correlation energy.[12] Because chemists are generally interested in energy differences which are very small relative to the total energy of the molecule, it is almost always necessary to recover a significant portion of the correlation energy to get useful results.[13]

Electron correlation arises from limitations inherent in the simple orbital model, which allows us to think about the properties of single electrons rather than the full multi-electron wavefunction. However, inter-electron interactions inherently couple the behaviour of all electrons, so producing this convenient single particle picture involves simplifying these interactions. Orbital-based models assume that the position of each electron does not depend on the specific positions of each other electron but rather their time-averaged positions.[14] This means that the position of each electron is less "correlated" with that of the other electrons than it would be in a true multi-

electron solution to the Schrödinger equation. Since the electrons do not experience direct Coulomb interactions, which serve to keep them apart more strongly than averaged interactions, there is a non-zero probability of two electrons being arbitrarily close.[1] In orbital-based models, multiple electrons are generally more stable occupying the same regions of space, and the electron density less diffuse than it should be.[15] Because this affects the dynamic movements of electrons relative to each other, this error in HF theory is often referred to as "dynamic correlation".

A second significant assumption in this simple picture of MOs is that there is a single "correct" electron configuration rather than a superposition of different configurations. This assumption is valid for many chemical systems at bonding distances, since in these systems the HOMO and LUMO have very large energy differences and the single lowest-energy electron configuration dominates the wavefunction. However, it breaks down when there are other configurations of a similar energy and the same symmetry as the lowest energy configuration; these configurations should be able to mix and form a wavefunction that is a superposition of multiple configurations.[16] In these cases where the mixing is significant, the wavefunction of the molecule cannot be modelled even qualitatively by any single determinant, and single determinant models frequently become unsuitable even as a starting point for more sophisticated models of the wavefunction.[17] In contrast to dynamic correlation, the error this introduces is generally referred to as "static correlation".[18]

1.2 Electron Correlation in H_2

Both classes of correlation can be demonstrated in the dissociation curve of the H_2 molecule. Figure 1.1 shows the binding energy curve of H_2 at different distances using three different wavefunctions. The single determinant wavefunction is an HF wavefunction - *i.e.* a single Slater determinant - constructed from a fully-optimized doubly occupied bonding MO. The two-determinant model expresses the wavefunction as a linear combination of this determinant and another constructed from a

fully-optimized antibonding MO. The full determinant model is a linear combination of all determinants corresponding to all possible electron configurations within the set of all available MOs; this exactly solves the Schrödinger equation to the limit allowed by the completeness of the atomic orbital basis set from which the MOs are constructed.[1] This allows us to compare different models of the wavefunction vs an essentially exact solution. Exact calculations are possible only for very small molecules, as the number of configurations scales exponentially with the number of orbitals.

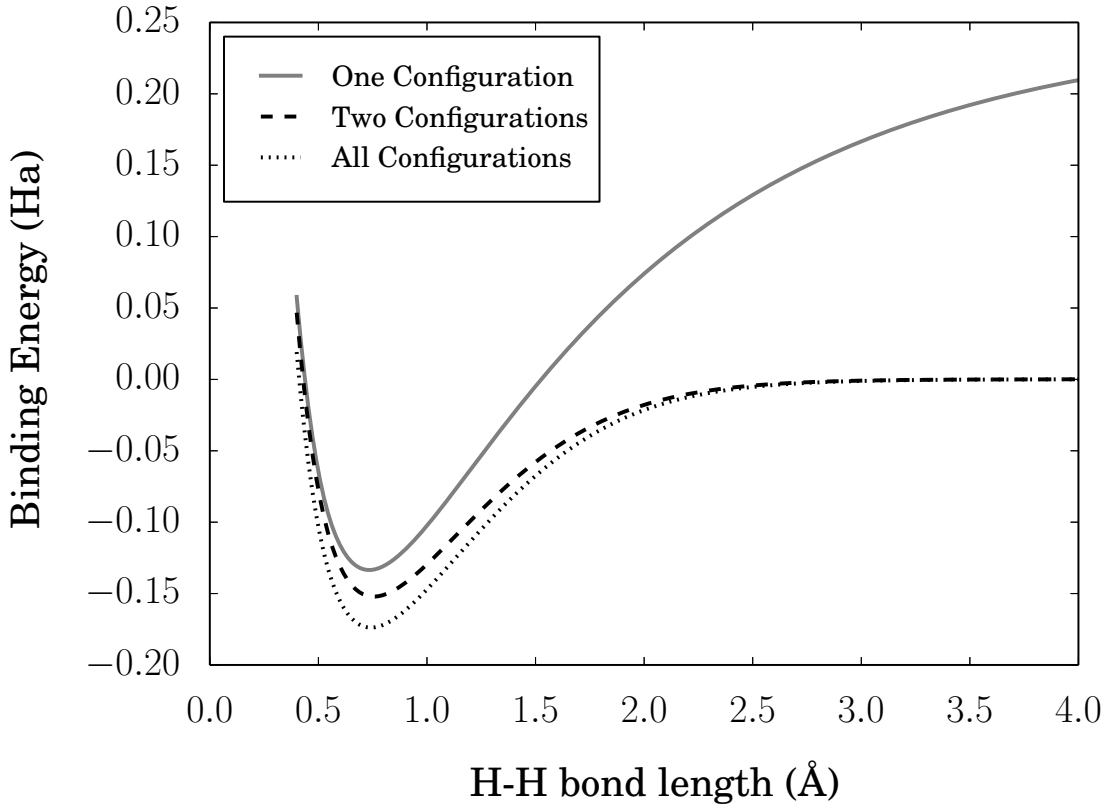

Figure 1.1: The dissociation curves of H_2 as calculated using one, two and full configuration models. At least the lowest energy and doubly-excited configurations are needed to model dissociation to independent atoms correctly, and many more are needed to fully account for the electron-electron interactions at bonding distances.

1.2.1 The Single Determinant Model

The single determinant (HF) model is approximately correct at bonding distance but has qualitatively incorrect behaviour at dissociation; instead of converging to the

energy of 2 H atoms the HF energy increases to a much larger value. At dissociation this wavefunction has 50% ionic character as of a H^+ H^- pair rather than correctly dissociating to two neutral atoms. In the single determinant wavefunction the positions of the electrons are independent but spatially equivalent. This is acceptable at bonding distances, when the electrons are delocalized over the whole molecule; but at dissociation results in the electrons being distributed statistically between the two nuclei, *i.e.* there is a 50% contribution from terms that involve two electrons localized onto the same atom, giving the wavefunction artifactual ionic character, and a 50% contribution from "covalent" terms that have one-electron associated with each atomic centre.

Another perspective on this problem comes from looking at the energies of the bonding and antibonding molecular orbitals. At equilibrium, these energies are very different, and so it is clearly energetically favourable to doubly occupy the bonding orbital only. However, at dissociation, the two energy levels become degenerate, so it is not clear which MO should be doubly occupied. In fact, neither configuration alone allows for qualitatively correct dissociation behaviour, and it seems logical that both determinants should contribute equally to the wavefunction if the underlying MOs are energetically equivalent. Since near degeneracy of the HOMO and LUMO is ubiquitous in systems with stretched or broken bonds, recovering static correlation is essential for predicting the energetics of bond making/breaking processes, *i.e.* modelling chemical reactivity and thermodynamics.[16]

1.2.2 The Two Determinant Model

By introducing a second determinant based on the doubly occupied anti-bonding orbital, we can include its contribution to the wavefunction. This two-determinant wavefunction has the properties we would expect from qualitative quantum theory, *i.e.* the closer the energy of the two configurations, the more they mix, and the more the anti-bonding determinant contributes to the final wavefunction. This means that the wavefunction has the conceptually correct behaviour of being dominated by the

bonding configuration at bonding distances - where this configuration is much lower in energy - and being an equal mixture of two configurations at dissociation, where they are degenerate. The doubly occupied anti-bonding orbital configuration interferes destructively with the part of the wavefunction that places the two electrons on the same atom. At infinite separation, the wavefunction is equal parts each of the configurations so the problematic parts of the wavefunction are completely cancelled, and when one electron is observed on one atom there is a 100% chance of the other electron being observed on the other atom. This means the ionic term in the wavefunction is removed and two-determinants are able to model dissociation correctly.

1.2.3 The Many Determinant Model

The two configuration wavefunction is conceptually correct, *i.e.* it is a combination of two near-degenerate configurations able to reproduce the correct shape of the dissociation curve and the correct energy at the dissociation limit. However, it still significantly underestimates the dissociation energy. This must be due to failing to recover the total energy at equilibrium, because the total energy at dissociation is known exactly, at twice the energy of an isolated hydrogen atom.

At equilibrium, the wavefunction is dominated by a single configuration, suggesting that the remaining correlation energy is due to the mean-field simplification of electron-electron interactions rather than the single-reference approximation. This suggestion is also supported by the fact that the two-configuration model converges to the same energy as the full-configuration model at dissociation, where the electron-electron interactions go to zero (for H_2). The two determinant model also lowers the energy at equilibrium compared to the one-determinant model, suggesting that even the second determinant introduced to solve the single-configuration problem also corrects partly for the mean-field approximation.

This example demonstrates two key features about the interrelation of these two

classes of electron correlation. Firstly, that representing the wavefunction as a sum of different configurations is a strategy that accounts for both forms of electron correlation and in fact the two forms are tightly interconnected.[19,20] Methods that are based on the idea of undoing the single-reference approximation will also improve the description of inter-electronic interactions over the mean-field approximation. Secondly, the error due to the qualitative failure of the one-determinant model at dissociation dwarfs the relatively small differences between the three models at equilibrium. This implies that in cases where a single-determinant model is qualitatively inappropriate, there is little point in trying to correct for the mean-field approximation without first solving the problems introduced by the single-determinant model. In other words, if the fundamental physics of the approximate wavefunction is incorrect, this can't be fixed by refining the details of electron-electron interactions.

1.2.4 Summary

The fact that the HF model fails even for the simple case of H_2 dissociation clearly illustrates the need for a better electronic structure model than the basic MO approach. To model conceptually correct wavefunctions for multi-configuration systems, we need to abandon the single Slater determinant picture of the wavefunction. Beyond this, achieving quantitative accuracy requires models that also correct for the mean-field approximation, even for single-reference systems.

In practice, all higher-level computational methods use HF theory as a starting point, then deploy various strategies of mixing in different electronic configurations.[21] These post-HF models can be separated into two categories. Single-reference models use one HF determinant as a "reference" for which the MOs are optimized, and that single set of MOs is used to construct the all the different determinants. In contrast, multi-reference models attempt to rigorously undo the single configuration approximation, by optimizing the MOs for multiple determinants, thereby avoiding treating the reference determinant differently to other determinants that contribute to the overall wavefunction.

1.3 Single-Reference Post-Hartree-Fock Methods

1.3.1 Perturbation Theory

A common method used in computational physics to solve many-body problems is perturbation theory (PT), where the solution to an insolvable problem is approximated as a solution to a solved problem plus a small correction or "perturbation". When perturbation theory is applied to HF solutions the perturbation can then be expressed as allowing mixing between the reference determinant and some other determinants. Each contribution is scaled by both the energy difference between the two configurations and their Coulomb-repulsion weighted overlap.[22] Perturbation theory provides a cheap method that accounts for some of the error associated with the mean-field approximation. A series of methods (the MPn series) have been developed which provide an increasingly sophisticated approximation to the post-HF part of the wavefunction by increasing the number of configurations allowed to mix with the reference. The most widely used of these is MP2, which includes only doubly-excited configurations relative to the reference.

MP2 calculations are among the cheapest post-HF corrections making this method a mainstay of computational chemistry - particularly for larger molecules.[23] However perturbation theory is non-variational *i.e.* increasing the order of the perturbation correction by allowing higher energy configurations to mix does not always result convergence to the exact energy. This can occur even in simple closed-shell systems like the neon atom and hydrogen fluoride.[23] It's commonly the case that MP2 gives better results for molecular properties than more costly MP3 calculations.[24]

Perturbation theory is an appropriate strategy for correcting a single reference determinant and capturing dynamic correlation. However, in multi-configuration cases the HF wavefunction is *by definition* not a good approximation to the true solution to the Schrödinger equation, *i.e.* the foundational assumption of perturbation theory doesn't hold. More generally, quantum mechanics provides no physical reason for treating a single configuration in any special or unique way - as perturbation

theory does for the reference HF determinant. Consequently, methods which treat configurations in a more equal way are likely to be more effective at modelling the contributions of multiple configurations to the wavefunction.

1.3.2 Configuration Interaction

The most obvious and direct way to model a superposition of configurations is using the configuration interaction (CI) method,[25] in which the CI wavefunction is explicitly constructed as sum of different configurations each with weights optimized to minimize the energy. In principle this is an ideal approach: it describes the overall wavefunction in a physically intuitive way, *i.e.* using a superposition of quantum states. It is also numerically robust and variational - increasing the number of configurations in the expansion will always decrease towards the exact energy. Including all possible electron configurations gives the exact solution to the Schrödinger equation within a given atomic orbital basis,[26] this is called full CI (FCI).

However, in practice, FCI is intractable for systems with more than approximately a dozen "active" or valence electrons.[27] To reduce the computational complexity, the CI expansion must be truncated. This is usually achieved by including all configurations with up to a certain number of electron excitations from the reference configuration *e.g.* a configuration interaction with singles and doubles (CISD) calculation would include all configurations that differ from the ground state by moving one or two electrons between MOs. Since only the double excitations mix directly with the ground state, this is often sufficient at equilibrium geometries.[28] Including higher degrees of excitation in the CI expansion gives a sequence of computational models defined by the maximum level of excitation (CIS, CISD, CISDT ect.); due to the variational principle these give increasingly accurate energies, and converge to the best energy possible for the give AOs at the limit of full CI.[29] However, if the wavefunction is multi-configurational large numbers of configurations can be required to correctly model it correctly. In multi-reference systems, this makes CI methods costly and also makes the physical interpretation of the wavefunction more

difficult.[30]

A fundamental problem with single-reference CI models is that a single determinant is treated differently to all the others. All determinants are equivalent during optimization of the CI expansion, but the underlying MOs are only optimized to best describe a single "reference" determinant. Because reliably converging the HF equations to any solution other than the global minimum is very difficult, the lowest energy, ground state configuration is usually taken as the reference.[31] The corollary is that MOs used in constructing excited state determinants are not optimized for that particular state.[32] Furthermore, the excited determinants are constructed using the unoccupied orbitals from the HF calculation, which do not contribute to the HF energy and therefore are unconstrained by the HF optimization process, often producing physically meaningless MOs.[31]

These two factors combine to mean that the higher energy determinants in the CI expansion are a poor description of the actual configurations,[33] and therefore a poor basis to represent the wavefunction. Consequently these CI expansions must rely on very large expansions to correct for the single-determinant approximation, rather than a smaller number of well-optimized physically meaningful determinants.

Another weakness of truncated CI methods is that they do not scale well to large systems due to a lack of both size consistency and extensivity. Size consistency is the ability of the methods to yield the same energy in a single calculation on two non-interacting systems as the sum of the energies of two separate calculations on each of those systems *i.e.* $E(A + B) = E(A) + E(B)$.[34] Size extensive methods return an energy that scales linearly with the number of electrons in the system.[35] Size consistency is particularly important in achieving correct dissociation curves since the system dissociates into two non-interacting fragments,[36] CI can give a qualitatively correct description of the system at dissociation, but usually recovers less correlation energy compared to separate calculations on the fragments.[37]

Full CI is both size consistent and extensive however truncated CI is neither. Trun-

cating the CI expansion means that we fail to correctly include all local excitations to the same degree *e.g.* a separate calculations on the two systems would include all double excitations on both systems at CID level while a calculation of the combined system would have to include up to quadruple excitations to include the case of simultaneous local double excitations.[37]

1.3.3 Coupled Cluster Theory

A similar approach to CI with many of the same advantages and disadvantages is the coupled cluster (CC) method,[38] which similarly constructs the wavefunction as combination of different determinants. CC however uses a different form of configuration expansion based on products of configurations - in constant the linear expansion used in CI.[39] However, unlike CI, the CC wavefunction also includes products of excitations which effectively introduces extra higher-level configurations. For example the CCSD wavefunction will contain quadruple excitations from the products two of double excitations.[40] Consequently, coupled cluster theory is size-consistent and size-extensive,[40] and the CC energy converges more quickly with respect to maximum excitation level than the CI energy.[41,42] Unlike CI, the coefficients of each of the determinants are not found by direct optimization but instead by iteratively solving a set of coupled equations.[43] The implementations of CC theory commonly used in computational chemistry are not variational below the full CI limit.[44,45]

CC is widely considered the best quality method for single configuration systems; in particular CC with single and double substitutions along with perturbative corrections for triples[46] - CCSD(T) - is commonly referred to as the "gold standard" in computational chemistry methods. Coupled cluster methods routinely yield relative energies to "chemical accuracy" (< 4 kJ/mol) but remain computationally tractable for systems with up to 20-30 atoms[47,48] However, coupled cluster models can fail for intrinsically multi-reference systems for the same reasons as CI.

1.4 Multi-Reference Post-Hartree Fock Methods

1.4.1 Multi-Configurational Self-Consistent Field

A solution to the deficiencies of the CI and CC methods is to optimize the orbitals used in the higher-energy determinants so that they are a better representation of higher-energy configurations. One approach to this is a multi-configuration self consistent field (MCSCF) calculation, which uses the same form of wavefunction as CI; however the MOs are optimized to minimize the energy of the entire multi-configuration wavefunction rather than just a single determinant. This is in contrast to CI where the MO coefficients are held constant after the HF reference determinant is found.[49] This avoids the large CI expansions required in CI like methods by ensuring that the MOs are variationally the best to represent the CI wavefunction. As a result, MCSCF wavefunctions can be much more concise - being qualitatively correct while using only determinants that have a physical reason to be in the expansion rather than being included to correct for poor MOs.[50,51] Since the determinants in MCSCF wavefunctions are truly treated equivalently, these are considered the best methods of calculations on multi-configuration systems.

However MCSCF calculations are very costly per configuration relative to other methods. This is because optimizing the MO and CI coefficients concurrently requires an iterative process with a CI calculation performed each iteration.[52] As a consequence, the set of determinants in a MCSCF calculation needs to be carefully selected to ensure the wavefunction captures the relevant physics without including expensive redundant determinants. A number of different schemes have been proposed to guide these decisions; however, these schemes still require prior knowledge of the nature of the states of interest, and are frequently only applicable to a subset of systems.[53–55] In general, no automated way of selecting determinants exists.[56]

Because of this problem, the most common form of MCSCF calculation is the complete active space SCF (CASSCF) method in which the CI expansion is composed of all rearrangements of the electrons with a set of active MOs, with the remaining

electrons and orbitals left unchanged.[56,57] This reduces the problem of choosing a set of determinants to one of choosing a set of relevant orbitals. This selection can be made using the same criteria discussed above in considering if a system is likely to be multi-configuration in character - unoccupied orbitals which have the same symmetry and similar energy to the HOMO are likely to be significant. CASSCF simplifies the process of designing MCSCF calculations, but choosing determinant sets that capture the relevant physics of the system without introducing unnecessary costs can still require significant domain knowledge and trial and error.[58]

Since the cost of orbital optimization limits the number of orbitals that can be included in the active space, it is common to apply multi-reference perturbation or CI methods to the MCSCF wavefunction to account for the remaining dynamic correlation.[59]

Although MCSCF provides a rigorous approach to orbital optimization, this comes at a significant computational cost, which, in turn, requires hand-picked determinant sets to construct useful wavefunctions in practical time frames. If the cost of optimizing orbitals for the higher energy determinants could be reduced, it would allow for smaller CI expansions without the need for such austerity in the determinant sets. In the interest of solving this problem, a number of methods have been developed to generate fully or partially optimized excited determinants in a cost effective way.

1.4.2 Spin-Flip Configuration Interaction

The first of these is the spin-flip CI (SFCI) method developed by the Krylov group.[60] Spin-flip CI is identical to regular CI except in its choice of reference determinant. In conventional CI, determinants are constructed from a basis of ground state HF orbitals. In SFCI, those orbitals are obtained from a higher spin multiplicity solution to the HF equations, before reversing the spin flip and rearranging the electrons within the higher multiplicity MOs.[61,62]

This ensures that all the relevant MOs are at least partially occupied in the HF step and therefore contribute to the energy and are optimized. Although these MOs are still only optimized for a single state they are - in principle - on average a better basis overall for the different determinants in the expansion. The spin-flip method has been demonstrated to be effective at modelling bi- and tri- radicals as well as bond breaking.[63]

However the MOs are now not optimized for the lowest-energy determinant and thus, in single-reference cases where the wavefunction is dominated by this determinant SFCI may give a worse representation of the wavefunction than simple HF. This effect is more pronounced for higher multiplicity references, where the electron configuration and orbitals differ more from the ground state. This means that for a SFCI calculation to be accurate in modelling processes such as bond dissociation, where the dominant configuration changes, it may still be necessary to use large CI expansions.

1.4.3 Non-Orthogonal Configuration Interaction

One way to avoid the problem of biasing the CI expansion towards a specific reference state is to use multiple sets of orbitals to construct each determinant with orbitals optimized for its specific configuration. This approach breaks the orthogonality between the configurations that most CI codes assume and rely upon. Consequently, such methods are called non-orthogonal configuration interaction (NOCI), and require additional code to account for the overlap between the different - no longer orthogonal - determinants.[64]

The key difference between different NOCI implementations is how the different determinants that form the basis of the NOCI expansion are obtained. They may be obtained from HF calculations with different spin multiplicities (spin-flip NOCI) or by finding non ground-state solutions to the HF equations.

The advantage of the spin-flip approach is its simplicity and numerical robustness.

However, a bias towards the ground state configuration remains, because it is the only configuration for which the spin multiplicity of the reference orbital set matches the spin multiplicity of the final determinant in the NOCI expansion. In other cases, higher spin multiplicity reference MOs are used as a basis for expanding determinants of lower spin multiplicity states.

Another approach would be to attempt to find the higher-energy solution to the HF equations corresponding to each configuration in the CI expansion. However, finding non-ground-state solutions to the HF equations is complicated by the overwhelming bias towards converging to the ground state in the general global optimization procedure. Some clever work-arounds have been proposed, including the Maximum Overlap Method, which constrains MO occupancy at each step during the HF optimization process based upon overlap with occupied orbitals from the previous iteration.[31] Another approach is "SCF Meta-Dynamics" which performs a series of HF calculations and adds terms to the HF equations to increase the energy of previously found configurations - making an excited determinant the new lowest energy solution.[65] Using HF solutions in this way will likely result in higher quality determinants, as the HF solutions represent the best single-reference approximation to the higher energy states. However, reliably finding these higher energy HF solutions is far from routine and convergence to any specified state of interest cannot always be guaranteed.[31,65]

1.4.4 Valence Optimized Coupled Cluster

A third approach is the valence-optimized-orbital coupled-cluster method with double excitations (VOO-CCD).[66] Analogous to MCSCF calculations, the molecular orbitals are optimized to minimize the energy of the CC wavefunction, within the valence active space. This yields a more compact wavefunction than the CASSCF wavefunction within the same valence active space. VOO-CCD is also variational, size-consistent, size-extensive and uniquely determined once the valence active space is algorithmically defined. In principle, these qualities make it an ideal multi-

reference method, but in practice the coupling between the MO coefficients and the determinant amplitudes makes the VOO-CCD optimization process numerically unstable. These numerical stability problems are pronounced enough that the section describing the VOO-CCD method in the Q-Chem quantum chemistry program manual feels the need to assure the user that "the program has nothing against you personally".[67] This means that in practice these methods don't provide a universally applicable alternative to CASSCF calculations.

1.4.5 Generalized Valence Bond Theory

A class of method that are very different in principle from any of the above are generalized valence bond (GVB) methods. Valence bond wavefunctions are constructed directly using combinations of orbitals localized on atoms or molecular fragments.[68] This makes them a more natural way of representing broken bonds or other multi-configuration systems with localized electrons *e.g.* diradicals.[69] However, they are less suitable for systems with a high degree of delocalization, *e.g.* benzene, where they predict a structure with alternating long and short C-C bonds.[70]

The GVB wavefunction is usually too expensive to calculate and two common simplifications are the perfect-pairing[71] and imperfect-pairing[72] models, which greatly simplify the spin part of GVB wavefunction by limiting the number of combinations of localized orbitals to those required to describe bond dissociation and constraining the rest of the orbitals to a Slater determinant-like core space. These methods, however, require starting orbitals that correspond to physically meaningful bonding and anti-bonding pairs, a process that is difficult to do in a general, reproducible, "black-box" way.[71,73]

1.5 Thesis Aim

The properties of methods that effectively account for the two classes of error are quite different.[74] Effective multi-configuration methods generally place great impor-

tance on optimizing orbitals while effective dynamic configuration methods generally rely on larger expansion series with lower quality orbitals since when all configurations are included all methods give equivalent results regardless of orbital quality. Generally, errors associated with the single determinant approximation can be addressed by correlating a small number of electrons in orbitals near the HOMO and LUMO - so long as those orbitals are well described - while all electrons contribute to the error associated with the mean-field approximation, necessitating larger CI expansions to account for dynamic correlation.[18]

Despite this, most post-HF methods attempt to solve both problems simultaneously. This is the approach taken by existing spin-flip and NOCI approaches: finding more optimized orbitals to account for static correlation, while *also* using large sets of configurations to capture dynamical correlation. We hypothesize that more efficient and accurate multi-reference methods can be found by separating out the two sources of error as much as possible, and using methods better suited to each class separately to improve upon the HF wavefunction. Much of the difficulty in modelling multi-configurational systems lies in finding a good, compact first-order approximation to the true wavefunction.

MCSCF methods can reliably model multi-configurational systems, but are frequently too expensive to be practical and are also critically reliant on the user defining a chemically-appropriate active space. We propose three criteria that an ideal multi-configuration method would fulfil:

1. The configuration expansion must be concise - this restrains the scaling of the method and helps ensure there is a clear chemical interpretation to the result. Achieving this requires the use of optimized orbitals tailored to each contributing configuration.

2. The configuration space must be determined automatically and algorithmically, without requiring user input, and must result in a model that yields continuous potential energy curves.

3. The optimization of the MOs and configuration series must be numerically stable to make the method practical without extensive trial and error.

None of the existing methods all three of these criteria. The aim of the present work is to develop and validate novel minimal-determinant multi-reference models for modelling static correlation that fulfil these criteria, and to explore their suitability as a basis for applying subsequent dynamic correlation corrections.

Chapter 2

SF-NOCI Theory

2.1 Model Construction

2.1.1 Existing NOCI Approaches

The previous chapter outlines how any model of the wavefunction must be capable of describing the inherently multi-configurational nature of many chemical systems for it to be widely applicable to the kind of problems chemists care about. A number of common methods based on multi-configuration wavefunctions fail to do this effectively because they use configurations with unoptimized orbitals. MCSCF methods solve this problem but at substantial computational cost and loss of generality.

An effective model for a first-order representation of the wavefunction must provide a cheaper way of optimizing orbitals and CI coefficients without requiring extensive user input to define the wavefunction and/or tweaking of calculation parameters to achieve convergence. This is the goal of NOCI methods, including spin flip NOCI.[75] Unlike in "regular" spin-flip based CI, spin-flip non-orthogonal CI uses the best available reference orbitals for any given excited state. For example, singly excited state determinants are constructed from a basis of triplet spin multiplicity ground state orbitals. Although these orbitals are not exactly optimized for the excited state of interest, they represent the closest possible approximation that can be found in

a routine, robust and numerically stable way.

However, existing applications of NOCI theory do not apply any restrictions to the configurations included in the CI expansion, beyond specifying a maximum excitation level; as in orthogonal CI. Instead, these models effectively just modify conventional CI expansions by switching out excited state determinants derived from ground state HF MOs for alternatives obtained from higher multiplicity reference calculations. This has the net effect of improving the convergence of the CI expansion but without isolating the configurations that give rise to the deficiencies of the single-reference approximation.[76]

2.1.2 An Alternative Approach

We propose a different route, following the extremely effective strategy from single-reference computational chemistry where a simple approximation (HF theory in the single-reference case) is used to find a simple but robust and well understood first order model of the wavefunction, that is then refined. This is effective both in terms of providing both useful heuristic models - as demonstrated by the ubiquity of the MO picture derived from HF theory - and useful strategies for quantitative calculations - such as PT, CI and CC based methods that use the HF wavefunction as a starting point.

A minimal-determinant multi-configuration model would provide a picture of the wavefunction composed of only physically meaningful components that map directly to common concepts well-understood by chemists - bonding and antibonding orbitals, dissociation to physically meaningful electronic configurations, insight into the correspondence between molecular orbital and valence bond models.

In this chapter, we propose a series of severely restricted active space and minimal-determinant SF-NOCI models to fulfil these requirements. Our approach is based on implementation of NOCI developed by the Head-Gordon group and uses the same mathematical machinery, but differs from existing CI implementations in its very

limited approach to configuration selection. The remainder of this section will lay out the ideas behind our proposed models, and their physical justification, while the second section will describe the mathematics required to implement those ideas.

2.1.3 Configuration Selection in CI

CI methods are generally classified by how they select the determinants used to expand the wavefunction. Usually in CI the virtual orbitals are unoptimized and the determinants a poor representation of the electronic state they supposedly model. Consequently, it is a waste of time to ascribe much physical significance to these determinants beyond their level of excitation. In orthogonal CI, only the double excitations mix directly with the ground configuration and consequently have the biggest impact on the wavefunction; however other excitation levels can mix with the double excitations and therefore indirectly contribute to improving the representation of the ground electronic state.[25] To model dissociation, the required excitation level is equal to the number of electrons in the bonds being broken, *i.e.* ensuring correct behaviour at all points along the dissociation curve. For example, modelling the full potential energy curve for a dissociating double bond requires quadruple excitations. Using this number of excitations allows the CI to fully span the space qualitatively relevant to the bond-breaking, regardless of the quality of the determinants, and therefore incorporates the multi-configurational nature of bond breaking by brute force. However, the very large number of quadruple excitations makes this very costly in general.

The poor quality of the excited state MOs frustrates attempts in CI to select significant configurations based on any physical understanding of the system beyond counting excitation levels. This contrasts with MCSCF methods where the MOs are fully optimized and great care is often placed in selecting the physically relevant configurations.[50,77] The more a CI method can optimize its orbitals to approximate MCSCF, the more it can rely on selecting only physically-relevant configurations to model multi-configuration wavefunctions, and the less very large brute force expan-

sions should be needed.

2.1.4 Spin-Flip Approach to Orbital Optimization

The goal of spin-flip CI is precisely to find reference determinants with better MOs for CI expansions. SFCI uses MOs found for a higher spin-state reference to provide a more balanced representation of the configurations, less biased towards the lowest energy configuration. However, as the number of spin flips increase, single-reference SFCI becomes increasingly biased *against* the lower energy configurations as they become increasingly different from the reference configuration - the biasing problem is only partially solved. In general, the spin-flip method gives access to a series of reference determinants, each deriving from a different spin multiplicity, and with different biases towards the core and virtual orbitals of the ground state. A key question that arises in single-reference SFCI is then which of these determinants will best describe the system at hand.

The only apparent way to completely avoid the biasing of the CI expansion towards a particular set of orbitals - short of performing a costly multi-reference optimization - is to allow the use of a different reference for each configuration. This means using a series of different HF calculations to provide customised MO sets for each excitation level, instead of solving the much harder problem of finding the best single set of MOs for the CI expansion. However, this comes at the cost of having to solve the non-orthogonal configuration interaction problem to determine how to best combine these non-orthogonal MO sets when forming the overall CI wavefunction. With this process, we are able to use orbitals derived from the multiplicity that is closest to the target configuration and thus - we conjecture - the best representation within the spin-flip approximation.

Finally, it remains to specify: (a) which spin multiplicities are used in generating HF reference orbital sets, and (b) which electronic excitations are performed, and which orbital sets are used to represent each of the determinants in the wavefunction.

2.1.5 Complete Active Space SF-NOCI

Within the SF-NOCI framework there are a number of different approaches to selecting which configurations to include. The simplest and most comprehensive theory is complete-active-space SF-NOCI (CAS-SF-NOCI), in which determinants are constructed by generating all possible electronic configurations within the active space created by spin-flipping. This active space is defined as the set of singly-occupied orbitals in the spin-flipped Hartree-Fock calculation. Determinants are constructed order by order, according to the spin-flip level.

For example, the initial reference determinant is simply the HF wavefunction of the original spin multiplicity. All systems considered in this thesis are singlets in their ground state. For simplicity, a singlet reference in all illustrative examples going forward, but this approach is trivially generalisable to other spin multiplicity reference states.

Next, all determinants that can be generated in the active MOs from a triplet ROHF calculation are included, as illustrated in Figure 2.1. This is approximately equivalent to a CASSCF(2,2) calculation with semi-optimized orbitals for the excited state determinants. However, it is important to note that these quasi-optimized orbitals may look quite different to fully-optimized CASSCF orbitals.

Like CASSCF(2,2), this CAS-SFS-NOCI approach is expected to be appropriate as a 0th order reference for modelling single bond dissociation and/or the lowest-lying singly-excited states, and corresponding electronic transition energies. However, for many chemical systems, multiple bond dissociation/reorganisation occurs, and transitions to higher-energy electronic states are of importance in molecular structure characterisation and determination of optoelectronic properties.

Figure 2.2 illustrates the next set of determinants, *i.e.* all those generated in the spin-flip-doubles active space. The resultant CAS-SFSD-NOCI method is approximately equivalent to CASSCF(4,4) albeit with a range of important distinctions:

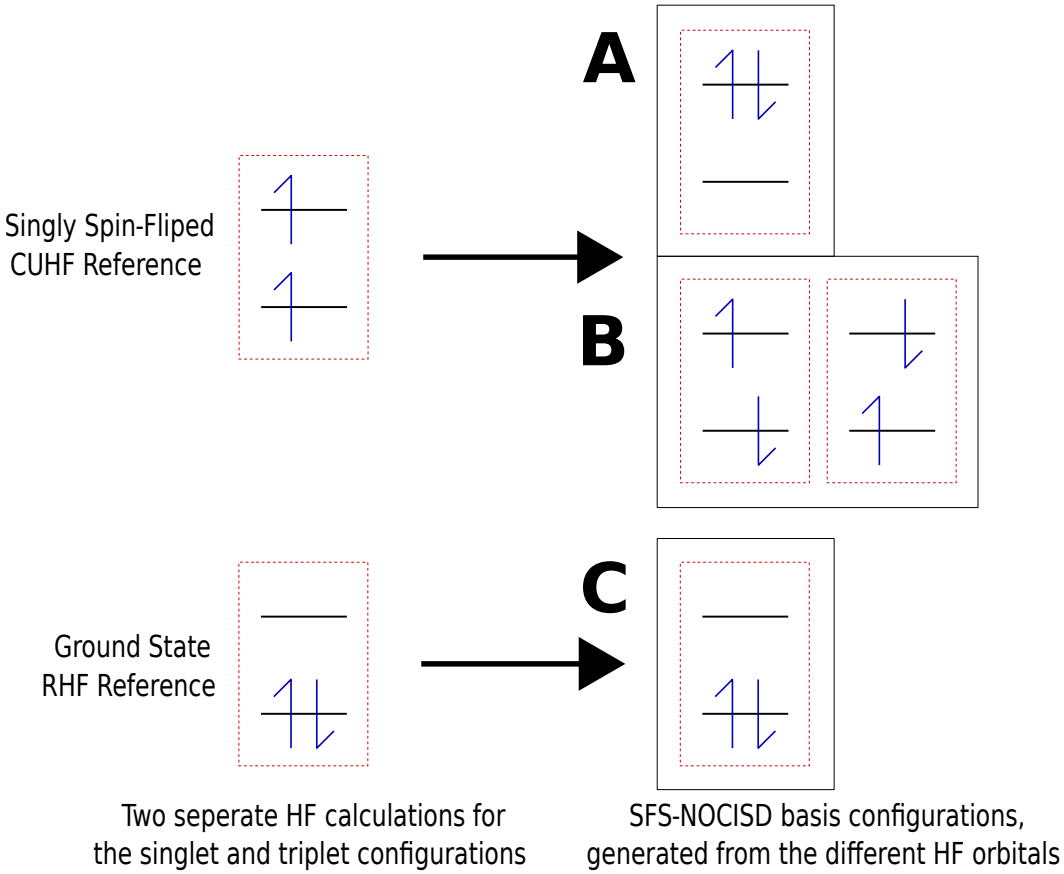

Figure 2.1: Generation of configurations for a SFS-NOCISD calculation. Triplet CUHF orbitals are used to form the singly-excited and doubly-exited configurations (groups A and B), singlet RHF orbitals are used for the ground-state configuration (group C).

- all sets of MOs used in forming the determinant basis are completely determined by the spin-flipping procedure i.e. they are generated automatically without any user input beyond specifying the maximum spin-flip level. This means that CAS-SF-NOCI methods are uniquely determined and can be used in a "black-box" manner.
- MOs are only semi-optimized for spin-flipped reference states rather than fully optimized for the target states of interest. However, this has the advantage of decoupling the orbital optimization and CI coefficient determination processes.
- different sets of MOs are used for expanding different determinants in the CI basis, *i.e.* the CI basis is non-orthogonal.

Overall, CAS-SF-NOCI models are robust, well-defined and applicable to a wide range of multi-reference problems.[64,75,76,78,79]

However, one disadvantage is that - like their CASSCF analogues - not all determinants within the active space necessarily have clear physical interpretations of the function they play in the wavefunction, *i.e.* CAS-SF-NOCI models are not minimal-determinant models for any particular problem of interest.

Therefore, the next two sections introduce new strategies for modelling specific multi-reference situations, aiming to include only those determinants required to ensure physically-correct behaviour, *i.e.* dissociation to appropriate atomic states and energies, inclusion of appropriate determinants for molecules that are multi-reference in character at equilibrium.

2.1.6 Flip-Reversing SF-NOCI

An alternative strategy for selecting configurations is performing the spin flip and then generating all configurations found by undoing the spin-flip without changing the occupancy of the MOs *i.e.* flipping spins but not performing any excitations or relaxations relative to the spin-flipped reference. This means that for a single spin-flipped reference, we would generate only the configurations in group B of Figure

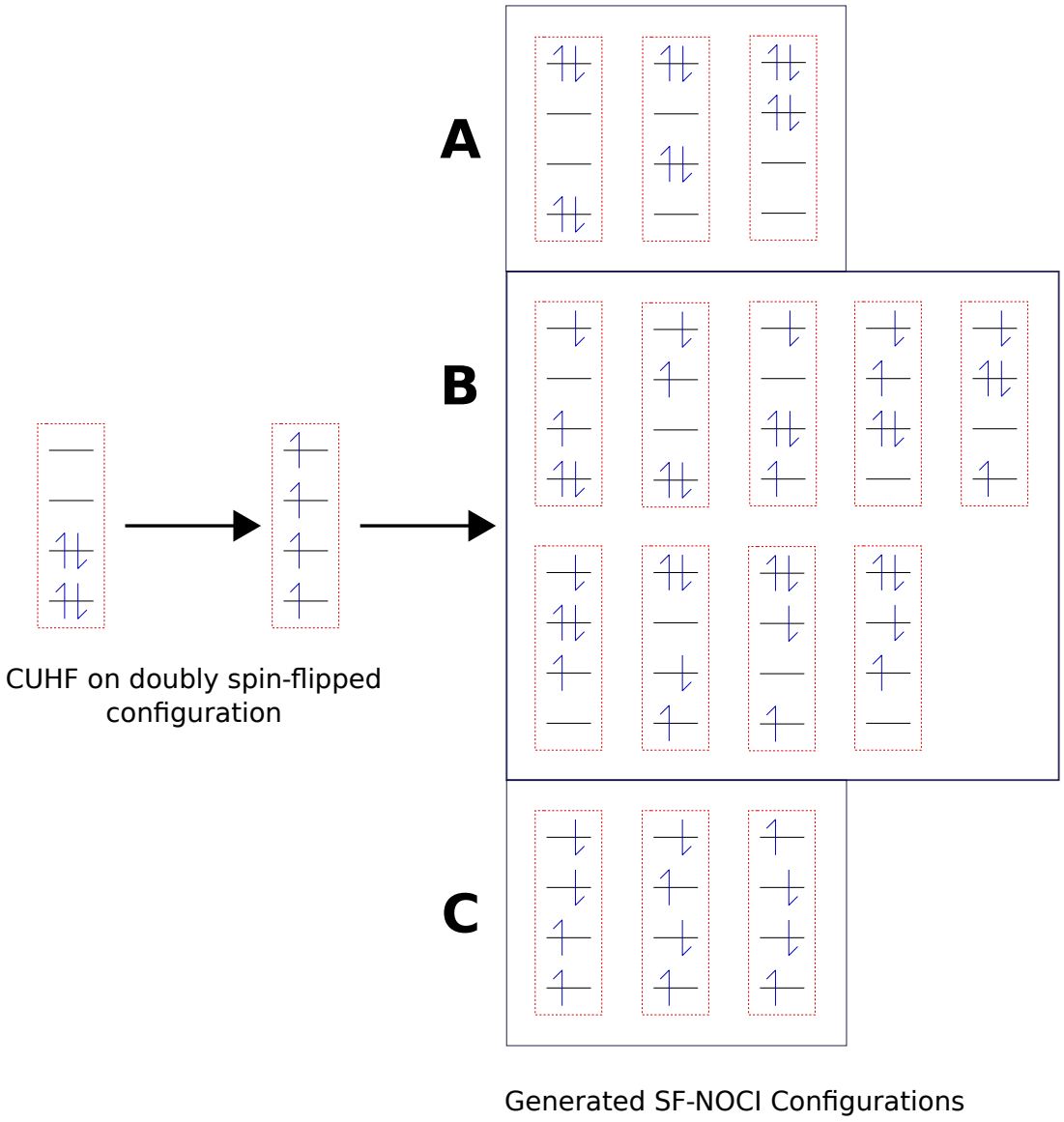

Figure 2.2: All possible SF-NOCI configurations generated from a single doubly spin-flipped determinant. Configurations with the 4th MO unoccupied are excluded as these would be constructed from references with fewer spin-flips, as shown in Figure 2.1 The configurations are organized into three groups: A) closed shell configurations, B) configurations with two unpaired electrons C) configurations with 4 unpaired electrons. Note that for the second two groups: B and C each configuration will have a corresponding spin-symmetric pair not shown here.

2.1 and for a doubly spin-flipped reference only group C of Figure 2.2. Unlike CAS-SF-NOCI, these wavefunctions depend only on the levels of spin-flip and not the degree of excitation within the spin-flipped state.

The simplest case of this would be a SFS-NOCIS, where the singlet and triplet references are used to generate 3 configurations, and corresponds to including the ground state reference augmented by the determinants shown in group B of Figure 2.1.

This approach is designed to efficiently model dissociation processes. Unpairing electrons by flipping the spin of one of them forces them to occupy spatially distinct molecular orbitals. At equilibrium this generally corresponds to bonding and antibonding orbitals, while at dissociation, high-spin ROHF calculations often localize electrons onto separate fragments.[80,81] Including all possible spin-symmetric flip-reversed determinants into the NOCI expansion ensures that the Pauli exclusion principle is observed and that the overall NOCI wavefunction has the appropriate spin symmetry, *i.e.* all eigenstates are eigenvalues of the $\hat{\mathbf{S}}^2$ operator.

FR-SFS-NOCI is therefore capable of modelling single bond dissociation, maintaining the correct spin symmetry throughout and smoothly transitioning from a wavefunction dominated by a doubly-occupied bonding orbital at equilibrium to an appropriate superposition of single-occupied atomic orbitals at dissociation. However, it is incapable of modelling systems that are inherently multi-configurational at equilibrium, due to the absence of the requisite doubly-excited determinant.

2.1.7 Perfect Pairing SF-NOCI

A third strategy with an even more restricted set of determinants is SF-NOCI with perfect pairing (PP-SF-NOCI). In the perfect pairing model, each electronic configuration differs from the ground state only by a double excitation of two paired electrons. This typically corresponds to excitation from a bonding orbital to the corresponding anti-bonding orbital (e.g. from the HOMO to the LUMO or the

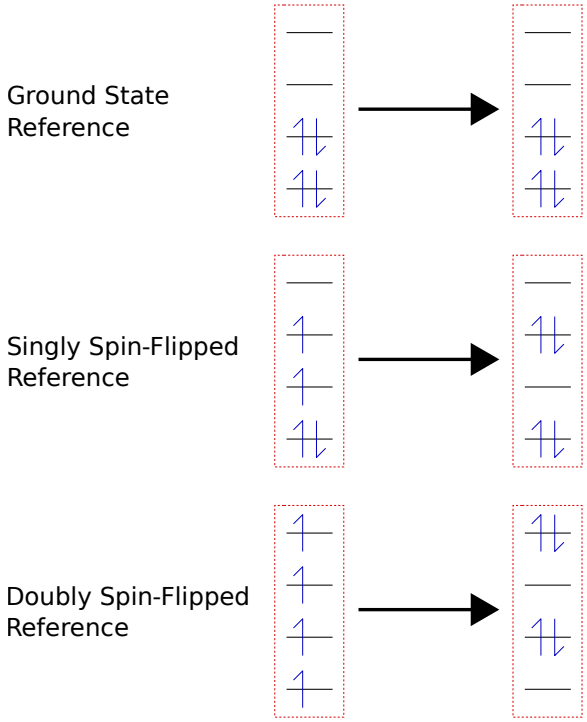

Figure 2.3: Generating determinants for a PP-SFSD-NOCI wavefunction. Each configuration includes only one paired double excitation from HOMO-n to LUMO-n. Consequently each spin-flipped reference determinant contributes only a single configuration to the CI.

HOMO-1 to the LUMO+1) as shown in figure 2.3.[82] However, ambiguities arise for systems with unoccupied atomic orbitals, in which two-electron $\sigma \rightarrow n$ and $\pi \rightarrow n$ excitations are possible, and lower in energy than the corresponding $\sigma \rightarrow \sigma^*$ and $\pi \rightarrow \pi^*$ configurations.

Nonetheless, simple perfect pairing methods have been used within valence-bond and coupled-cluster frameworks as a cheap way to account for the leading terms of correlation energy and found to be effective for diradical and bond breaking, provided that the MO coefficients are optimized concurrently with CI/CC coefficients.[83]

However, the key advantage of all SF-NOCI models proposed here is the fact that they decouple molecular orbital coefficient optimization from CI coefficient optimization. This is designed to both improve numerical stability and decrease computational cost. But this does come at a cost - PP-SF-NOCI models are not well suited to describing bond dissociation because they neither contain terms that allow for electron localization (like FR-SF-NOCI models) nor contain determinants with

fully-optimized antibonding orbitals, as required to accurately reconstruct atomic states as a linear combination of states involving doubly-occupied bonding and antibonding orbitals.

On the other hand, PP-SF-NOCI is ideally suited to describing systems that are strongly multi-reference at equilibrium, whose contributing states differ from the ground state configuration by excitation of pairs of electrons from occupied to unoccupied (not necessarily antibonding) orbitals.

2.1.8 Summary of SF-NOCI Strategies

The computational efficiency of SF-NOCI is a function of the number of determinants in the CI expansion basis. For each reference set of orbitals, this number can be quantified in terms of the spin-flip order: $F = 1$ (single spin-flip, triplet reference state), $F = 2$ (double spin-flip, pentet reference state), *etc*, and the maximum excitation order within the spin-flip active space, N (note that N must be $\leq 2F$)

For any given SF-NOCI method, computational scaling will be dominated by the highest order spin-flip level and excitation level, combined. The total number of determinants is computed as the sum of each set of determinants generated at each spin-flip order, where $\binom{n}{k}$ is the binomial coefficient $\frac{n!}{k!(n-k)!}$, and "# Configurations per reference" is the number of NOCI configurations generated from a single spin-flipped reference configuration.

Complete-Active-Space SF-NOCI

- **Strategy:** All possible arrangements of active electrons within the spin-flip optimized space.
- **Effective for:** A wide range of geometries in both excited and ground states.
- **# Configurations per Reference:** $\binom{2F}{N}^2 - \binom{2F-1}{N}^2$

Spin-Reversing SF-NOCI

- **Strategy:** All possible spin-flips back to the target spin multiplicity without changing spatial orbitals.

- **Effective for:** Accurately modelling molecular dissociation.
- **# Configurations per Reference:** $2\binom{2F}{F}$

Perfect Pairing Space SF-NOCI

- **Strategy:** Paired double excitations between the HOMO-n and LUMO+n.
- **Effective for:** Strongly correlated systems in the ground state and at bonding geometries.
- **# Configurations per Reference:** 1

2.2 Mathematical Basis

2.2.1 The Hartree-Fock Method

SF-NOCI calculations rely on the HF method to generate MOs for the later NOCI. HF theory generates the determinants that form the basis of the NOCI wavefunction, and it is necessary to define all quantities computed during the HF process before proceeding to define the equations required to implement NOCI. Therefore, a brief overview of Hartree-Fock theory is presented below. Atomic units will be used throughout.

The HF equation has the same form as the Schrödinger equation, however the electronic Hamiltonian is replaced by a Fock operator $\hat{\mathbf{F}}$ that approximates the electronic energy in terms of single electron orbitals.[84] The eigenvalues of the Fock operator are the MOs ϕ and the corresponding eigenvalues E are the negative of the ionization energy of an electron in that MO (Koopman's theorem).[85]

$$\hat{\mathbf{F}}(1) = \hat{\mathbf{h}}(1) + \sum_{b \neq a} \left[\hat{\mathbf{J}}_b(1) - \hat{\mathbf{K}}_b(1)\right] \tag{2.1}$$

$$\hat{\mathbf{F}}\phi = \mathrm{E}\phi \tag{2.2}$$

The Fock operator partitions the energy of each MO into three different classes, one for each of the three terms in equation 2.1. The first is the core Hamiltonian energy

i.e. the energy that depends on only single electrons, as shown in equation 2.3. This is further subdivided into electron kinetic energy (the first term of equation 2.3) and the electron-nuclei attraction (the second term). Here ∇ is the gradient operator in the space coordinates of electron 1, Z_A is the nuclear charge of atom A and r_{1A} the distance between the electron and atom A.

$$\hat{\mathbf{h}}(1) = -\frac{1}{2}\nabla_1^2 - \sum_A \frac{Z_A}{r_{1A}} \tag{2.3}$$

The two remaining parts of the Fock operator are the Coulomb $\hat{\mathbf{J}}$ and exchange operators $\hat{\mathbf{K}}$, which quantify the electron-electron interactions; the latter being a purely quantum mechanical effect arising from the requirement that the wavefunction be antisymmetric to electron interchange. We define these operators by their operation on a single spin orbital $\chi_a(1)$.

$$\hat{\mathbf{J}}_b(1)\chi_a(1) = \Big[\int \chi_b^*(2) r_{12}^{-1} \chi_b(2) \mathrm{d}\mathbf{x}_2\Big]\chi_a(1) \tag{2.4}$$

$$\hat{\mathbf{K}}_b(1)\chi_a(1) = \Big[\int \chi_b^*(2) r_{12}^{-1} \chi_a(2) \mathrm{d}\mathbf{x}_2\Big]\chi_b(1) \tag{2.5}$$

The integral over all coordinates of the second electron in these operators is the mathematical realization of the mean-field approximation, which follows from expressing the Fock operator as a function of the coordinates of a single electron.

Basis Sets

Solving the HF equations using a computer requires us to reduce the HF integro-differential equation to a matrix eigenvalue problem by selecting a set of "basis functions" to represent the orbitals. This is analogous to qualitative MO theory where hydrogen-atom-like AOs are used to build up the MOs. Virtually all molecular quantum chemistry codes use Gaussian functions (equation 2.6) centred on the nuclei for this purpose, as they strike a balance between resembling the solutions to the Schrödinger equation for the H atom, while also allowing the integrals required to evaluate the electron-electron interactions be calculated efficiently.[86] To improve

the quality of the basis functions, multiple Gaussians are combined with fixed coefficients c into a single contracted Gaussian type orbital (CGTO) (equation 2.7. This represented by the ϕ_{ijk}, where i, j and k denote the quanta of angular momentum along each axis.

$$G_{ijk}(r;a) = Nx^i y^j z^k e^{ar^2} \tag{2.6}$$

$$\phi_{ijk}(r;\hat{a},\hat{c},N) = \sum_n^N c_n G_{ijk}(r;a_n) \tag{2.7}$$

Where N is a normalization factor, c_n are the contraction coefficients which determine how much each Gaussian function contributes to the CGTO; the exponents a_n determine how diffuse each Gaussian is. Part of the setup of a computational chemistry calculation is the choice of the basis set - a pre-optimized set of CGTOs with fixed values of the c_n and a_n for each atom. Generally higher quality basis sets contain more CGTOs per occupied AO, in some cases specialized basis sets are required *e.g.* excited state calculations may require basis sets with unusually diffuse orbitals (small values of a_n) to properly model more diffuse electronic states.

The parameters i, j and k represent the components of the orbital angular momentum (L) along each axis; a CGTO will contain only Gaussian functions of the same angular momentum. Calculations will include all permutations of i, j and k that sum to the total angular momentum. This means that for the form used in equation 2.6, there will be 6 different d type functions ($d_{x^2}, d_{y^2}, d_{z^2}, d_{xy}, d_{xz}, d_{yz}$) rather than the expected 5, and 10 f functions rather than the expected 7 etc. In these cases, the orbitals can be transformed into a set of the expected number of orbitals plus an additional set of L-2 orbitals[87] e.g. an additional s type orbital for each set of d functions, and an additional three p orbitals for each f set.

The parameters $\{c_n\}$ and $\{a_n\}$ are optimized ahead of time for a set of test molecules and are held constant during the HF calculation. We can then apply the ubiquitous technique in quantum mechanics of representing an operator in a particular basis as a matrix (e.g. equation 2.8 for the Fock matrix in the basis of atomic orbitals). This

yields the Roothaan equations which reformulate the HF equation as a generalized eigenvalue problem, where each eigenvector in **C** contains the molecular orbital coefficients for a single MO.

$$F_{ij} = \left\langle \phi_i^* \middle| \hat{\mathbf{F}} \middle| \phi_j \right\rangle \tag{2.8}$$

$$\mathbf{FC} = \mathbf{SCE} \tag{2.9}$$

$$\psi_j^{MO} = \sum_i C_{ij}\phi_i \tag{2.10}$$

S is the matrix of basis function overlaps (in this case the overlaps of the CGTOs) and **C** is the matrix of molecular orbital coefficients. As equation 2.10 shows, each column of the MO matrix **C** represents a single MO, and each row represents a single CGTO *i.e.* C_{ij} is the contribution of the i^{th} CGTO to the j^{th} MO.

The Self-Consistent Field Procedure

HF calculations are concerned with finding the values for $\mathbf{C}_{ij}$ that minimize the energy of a single Slater determinant, Φ^{HF}.

$$\Phi^{\text{HF}} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_i(\mathbf{x}_1) & \chi_j(\mathbf{x}_2) & \cdots & \chi_k(\mathbf{x}_1) \\ \chi_i(\mathbf{x}_2) & \chi_j(\mathbf{x}_2) & \cdots & \chi_k(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_i(\mathbf{x}_N) & \chi_j(\mathbf{x}_N) & \cdots & \chi_k(\mathbf{x}_N) \end{vmatrix} \tag{2.11}$$

Where the $\chi_i(\mathbf{x}_n)$ is the i^{th} MO occupied by the n^{th} electron. Equation 2.12 shows the calculation of the electronic energy in the HF procedure.

$$E^{HF} = \frac{1}{2}\sum_i \sum_j P_{ij}F_{ij} \tag{2.12}$$

$$P_{ij} = 2\sum_a^{n/2} C_{ia}C_{aj} \tag{2.13}$$

P is the density matrix, which represents the electron density contained in the overlaps of each pair of the CGTOs. It is a key intermediate value in the calculation

of the Fock matrix from the MOs in most codes. The factor of 2 in the expression for **P** accounts for both the alpha and beta electrons.

Calculating the MO coefficients requires finding the Fock matrix, which itself depends on the MO coefficients making the HF equations non-linear and requiring an iterative solution process. The simplest such process is the Roothaan algorithm which uses simple fixed point iteration to find a solution. A guess set of MOs are used to calculate an initial Fock matrix; this is then diagonalized to generate a new set of MO coefficients. The new set of MO are then used to find a new Fock matrix and the process repeated until a fixed point is found and the MOs stop changing.

Since each step of the SCF procedure results in a number of MOs equal to the number of CGTOs - which is usually much larger than the number of electrons - only a subset of the MOs are "occupied" and used to calculate the next Fock matrix. These are usually chosen to be the MOs with the lowest eigenvalues. This choice gives rise to the division between the optimized occupied orbitals and the optimized virtual orbitals previously discussed.

In practice, the Roothaan algorithm often converges slowly or diverges, so a large number of different algorithms have been developed to refine or replace it; however, they all follow the same iterative logic of starting with guess MOs and refining them until self consistence is reached. Our implementation of SCF uses the Direct Inversion of the Iterative Subspace (DIIS), which uses information from previous iterations to improve the quality of the Fock matrix each Roothaan step.[88,89]

Constrained Unrestricted Hartree-Fock

The above discussion ignores the distinction between alpha and beta electrons and their orbitals. In the simplest formulation of HF theory - closed-shell restricted Hartree-Fock (RHF) - we assume that each MO is doubly occupied or equivalently that the spatial orbitals for the alpha and beta electrons are the same, and that there are the same number of "up-spin" and "down-spin" electrons. This assump-

tion obviously fails to account for open-shell systems such as doublets or excited states where the alpha and beta electron configurations are different - it also fails in describing dissociation, where the two bonding electrons dissociate onto different atoms, as discussed in the introduction.

Removing the restriction that the alpha and beta orbitals be identical gives Unrestricted HF (UHF).[90] This essentially involves performing two separate HF calculations, each with their own density, Fock and MO matrices but interacting via the Coulomb part of the Fock operator, resulting in paired UHF equations shown in equation 2.15 where $\mathbf{F}^{\alpha}(\mathbf{C}^{\alpha}, \mathbf{C}^{\beta})$ denotes that $\mathbf{F}^{\alpha}$ depends on both sets of MOs. This can result in wavefunctions which are not eigenfunctions of the spin squared operator but a linear combination of different spin eigenstates *i.e.* other states of different spin multiplicity are allowed to mix with the wavefunction, an effect called "spin contamination". At bonding distances in closed-shell systems, the UHF wavefunction usually reduces to the RHF wavefunction, which is a pure spin eigenstate, and at dissociation the different spin states contributing to the UHF wavefunction become energetically degenerate. Consequently UHF gives correct energies at bonding distances and dissociation (and hence correct dissociation energies) but unphysically high energies in the intermediate region[91]

$$\mathbf{F}^{\alpha}(\mathbf{C}^{\alpha}, \mathbf{C}^{\beta})\mathbf{C}^{\alpha} = \mathbf{S}\mathbf{C}^{\alpha}\mathbf{E}^{\alpha} \tag{2.14}$$

$$\mathbf{F}^{\beta}(\mathbf{C}^{\alpha}, \mathbf{C}^{\beta})\mathbf{C}^{\beta} = \mathbf{S}\mathbf{C}^{\beta}\mathbf{E}^{\beta} \tag{2.15}$$

A third approach called Restricted Open Shell HF (ROHF) allows the singly-occupied orbitals to differ but restrict the doubly-occupied to be the same regardless of electron spin.[92] This gives wavefunctions which are eigenfunctions of $\hat{\mathbf{S}}^2$ and hence do not suffer from spin contamination. However there is no unique form of the Fock matrix in ROHF theory and therefore no unique MOs or MO energies complicating both physical interpretation of the results and their use in post-HF methods.[93] The most common form of the ROHF Fock matrix generates the "canonical" HF orbitals,

which obey Koopman's theorem.[94]

In our calculations, we use a particular implementation of ROHF known as constrained unrestricted HF (CUHF). CUHF was chosen as it is simpler to implement than other ROHF constraint schemes but still yields canonical ROHF orbitals,[80,81] and - in our experience - displays very good SCF convergence behaviour. Using symmetry-broken UHF orbitals should be avoided in post-HF methods since the spin-symmetry breaking means that the spin contamination carries through to the post-HF wavefunction.[95]

In the CUHF process, the alpha and beta density and Fock matrices are first calculated as normal in UHF. The density matrices $\mathbf{P}$ and the AO overlap matrix $\mathbf{S}$ are then used to calculate the UHF natural orbitals $\mathbf{C}^{\mathrm{NO}}$.

$$\mathbf{P}^{\mathrm{NO}} = \mathbf{S}^{\frac{1}{2}} \frac{\mathbf{P}^{\alpha} + \mathbf{P}^{\beta}}{2} \mathbf{S}^{\frac{1}{2}} \tag{2.16}$$

$$\mathbf{P}^{\mathrm{NO}}(\mathbf{S}^{\frac{1}{2}}\mathbf{C}^{\mathrm{NO}}) = (\mathbf{S}^{\frac{1}{2}}\mathbf{C}^{\mathrm{NO}})\sigma \tag{2.17}$$

In these equations the superscripts denote whether the matrix is represented in the MO or NO basis.

We now split the orbitals into core, active and virtual spaces. The core space includes all doubly occupied orbitals, the active space all singly occupied orbitals, and the virtual space the remaining unoccupied orbitals. We fix these at the start of the calculation based on the number of electrons in the system and its specified multiplicity.

To constrain the MOs, we need to find the differences between the Fock matrices in the NO basis $\mathbf{\Delta}^{\mathrm{NO}}$. In particular, we need the differences in the terms that describe the interaction between the core and valence orbitals - the λ matrix which is zero except in the valence-core and core-valence blocks. The difference matrices

are computed as follows.

$$\mathbf{\Delta}^{\mathrm{MO}} = \frac{\mathbf{F}^{\alpha} - \mathbf{F}^{\beta}}{2} \tag{2.18}$$

$$\mathbf{\Delta}^{\mathrm{NO}} = (\mathbf{C}^{\mathrm{NO}})^{\mathrm{T}} \mathbf{\Delta}^{\mathrm{MO}} \mathbf{C}^{\mathrm{NO}} \tag{2.19}$$

$$\lambda^{\mathrm{NO}} = \begin{array}{c} \\ c \\ a \\ v \end{array} \begin{array}{c} \begin{array}{ccc} c & a & v \end{array} \\ \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{\Delta}^{\mathrm{NO}}_{cv} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{\Delta}^{\mathrm{NO}}_{vc} & \mathbf{0} & \mathbf{0} \end{pmatrix} \end{array} \tag{2.20}$$

$$\lambda^{\mathrm{MO}} = ((\mathbf{C}^{\mathrm{NO}})^{-1})^{\mathrm{T}} \lambda^{\mathrm{NO}} (\mathbf{C}^{\mathrm{NO}})^{-1} \tag{2.21}$$

We then use the λ^{MO} difference matrix and the two UHF Fock matrices to get the CUHF Fock matrices $\widetilde{\mathbf{F}}^{\alpha}$ and $\widetilde{\mathbf{F}}^{\beta}$, which can then be diagonalized as usual to get the next sets of CUHF MOs and the process repeated until convergence is reached.

$$\widetilde{\mathbf{F}}^{\alpha} = \mathbf{F}^{\alpha} + \lambda^{\mathrm{MO}} \tag{2.22}$$

$$\widetilde{\mathbf{F}}^{\beta} = \mathbf{F}^{\beta} - \lambda^{\mathrm{MO}} \tag{2.23}$$

2.2.2 Configuration Interaction Theory

CI theory provides a framework for combining the determinants constructed from a basis of HF MOs to create a better description of the wavefunction. CI theory works on the same basic variational principle as HF theory; it constructs the wavefunction as a linear combination of basis states (each described by a Slater determinant) and then optimizes the contribution of each of those states to the overall energy of the wavefunction. The key difference is that while CI wavefunctions are constructed from a set of fixed Slater determinants, HF wavefunctions correspond to a single determinant that is optimized by adjusting the linear combinations of atomic orbitals that form the underlying set of molecular orbitals in order to minimize the energy of that determinant.

Nonetheless, there is a clear correspondence between the form of the equations that must be solved to find the CI wavefunction and the Roothaan equations that define the HF procedure:[25]

$$\mathbf{HA} = \mathbf{ESA} \tag{2.24}$$

$$H_{ij} = \langle \psi_i | \hat{H} | \psi_j \rangle \tag{2.25}$$

This is essentially just the Schrödinger equation with the Hamiltonian operator $\hat{H}$ represented as a matrix in the basis of the determinants. The normalized eigenvectors **A** of the Hamiltonian matrix are the contributions of the determinants to each state and the eigenvalues **E** the electronic energies of each state. The matrix **S** is the matrix of determinant overlaps, analogous to the matrix of basis function overlaps in the Roothaan equations; in the case of orthogonal CI it is the identity matrix. Unlike in HF theory we are not attempting to optimize the basis itself (in this case the determinants of the CI expansion) - only the contributions of each determinant to the wavefunction, consequently we only need to solve this equation once (in contrast to MCSCF where the determinants are optimized, requiring an iterative process which solves the CI equations multiple times).

In conventional CI, a single HF calculation is performed and the determinants formed by permuting the columns of the MO matrix to simulate excitations from the occupied to virtual orbitals. Since the MOs are mutually orthogonal (as the Fock matrix is Hermitian) this means that the overlap matrix **S** is the identity matrix. The elements of the CI Hamiltonian can then be calculated using the Slater-Condon rules, which assume orthogonal orbitals.[96]

$$
\begin{aligned}
\langle\Phi|\hat{H}|\Phi\rangle &= \frac{1}{2}\sum_{i,j}\langle ij||ij\rangle + \sum_{i} h_{ii} \\
\langle\Phi|\hat{H}|\Phi_a^r\rangle &= \sum_{i}\langle ai||ri\rangle + h_{ii} \\
\langle\Phi|\hat{H}|\Phi_{ab}^{rs}\rangle &= \langle ab||rs\rangle \\
\langle\Phi|\hat{H}|\Phi_{abc}^{rsq}\rangle &= 0
\end{aligned}
\tag{2.26}
$$

$$
\begin{aligned}
\langle ij||kl\rangle = &\int \chi_i^*(1)\chi_j^*(2)\, r_{12}^{-1}\, \chi_k^*(1)\chi_l^*(2)\mathrm{d}\mathbf{x}_1\mathrm{d}\mathbf{x}_2 - \\
&\int \chi_i^*(1)\chi_j^*(2)\, r_{12}^{-1}\, \chi_l^*(1)\chi_k^*(2)\mathrm{d}\mathbf{x}_1\mathrm{d}\mathbf{x}_2
\end{aligned}
\tag{2.27}
$$

The notation Φ_{ab}^{rs} denotes a determinant where electrons have been excited from the MOs χ_a and χ_b into orbitals χ_r and χ_s relative to the reference Φ. The matrix $\mathbf{h}$ is the core Hamiltonian matrix defined in equation 2.3.

2.2.3 NOCI Theory

The use of a single universal set of orthogonal orbitals in conventional CI and SFCI simplifies the calculation of the elements of the CI Hamiltonian matrix. With the orthonormal determinants constructed from this orbital set, CI matrix elements can be calculated using the Slater-Condon rules.[96] Pairs of determinants can be easily classified based on the number of excitations separating the two electron configurations, with all differing orbitals being orthogonal to those of the other set *i.e.* a triply excited configuration will have three orbitals orthogonal to the ground state.

In SF-NOCI, we use a number of single-reference calculations to find our determinants rather than just one, this gives MOs that are a better basis for the higher energy configurations in a much simpler way than MCSCF.[76,79] However, it means sacrificing the mutual orthogonality of the determinants and introducing more complexity into calculation CI Hamiltonian matrix elements.

We have significant freedom to transform the HF orbitals to mitigate the complexity introduced by non-orthogonality. A well known property of the HF wavefunction is that it is invariant to mixing within the occupied orbitals,[4] *i.e.* so long as we restrict our transformations to unitary rotations within the occupied columns of the MO matrix, we can modify the MOs without changing the physical properties of the wavefunction. One approach to evaluating integrals over non-orthogonal orbitals is to transform the orbitals of each pair of determinants such that they are biorthogonal. This means that each orbital overlaps with, at most, one of the orbitals in the other determinant.[4]

$$\langle \phi_i | \widetilde{\phi}_j \rangle = s_i \delta_{ij} \tag{2.28}$$

Here δ_{ij} is the Kronecker delta and s_i is the overlap between the two orbitals $0 \leq s_i \leq 1$. This is a generalization of the single-reference fully-orthogonal case, which is identical except that $s_i = 0$ or 1. Therefore, the MO integrals required to form the Hamiltonian matrix can be calculated using a modified form of the Slater-Condon rules, in which each part of the integral between two determinants is scaled by the overlap between the appropriate biorthogonalized pair of MOs.

We use the formulation of non-orthogonal CI developed by the Head-Gordon group.[64,75,76] The relevant source code from our "pychem" software is contained in Appendix A. For each determinant pair, we use the Löwdin pairing theorem[4] to biorthogonalize the orbitals using a singular value decomposition.

$$\mathbf{U}, \sigma, \mathbf{V}^T = (\widetilde{\mathbf{C}}^{\mathrm{HF}})^T \mathbf{S} \mathbf{C}^{\mathrm{HF}} \tag{2.29}$$

$$\mathbf{C} = \mathbf{C}^{\mathrm{HF}} \mathbf{U} \tag{2.30}$$

$$\widetilde{\mathbf{C}} = (\widetilde{\mathbf{C}}^{\mathrm{HF}}) \mathbf{V} \tag{2.31}$$

Here $\mathbf{C}^{\mathrm{HF}}$ and $\mathbf{C}$ represent the pre- and post-biorthogonalization MO coefficients respectively, and $\mathbf{S}$ represents the atomic orbital overlap matrix. Note that here

$\mathbf{C}^{\mathrm{HF}}$ is truncated to include only the columns of the MO matrix representing the occupied orbitals; if the virtual orbital coefficients are included, the biorthogonalizing transform will mix the occupied and virtual orbitals, resulting in configurations that are physically different from the pre-biorthogonalized determinants.

The diagonal elements of σ, $\{s_i\}$ give the overlaps between each pair of MOs, the product of the these is S - the total overlap between the two determinants. We also calculate the reduced overlap S^{red}, which ignores the orthogonal MO pairs and hence will always be non-zero.

$$S = \det(\mathbf{V})\det(\mathbf{U}) \prod_{s_i} s_i \tag{2.32}$$

$$S^{\mathrm{red}} = \det(\mathbf{V})\det(\mathbf{U}) \prod_{s_i \neq 0} s_i \tag{2.33}$$

Here $\det(\mathbf{V})$ and $\det(\mathbf{U})$ are the matrix determinants of the two SVD transform matrices. Since $\mathbf{V}$ and $\mathbf{U}$ are unitary, their determinants will be ± 1 and represent the relative orbital phase, which is otherwise lost in the SVD procedure. With these orbitals, we can use the Slater-Condon rules generalized to biorthogonal MOs to calculate the electronic part of the CI Hamiltonian.

$$\begin{aligned}
\langle \Phi | \hat{H}_{\mathrm{elec}} | \widetilde{\Phi} \rangle &= \frac{1}{2} \sum_{i,j} \frac{\langle ij||ij \rangle}{s_i s_j} + \sum_i \frac{h_{ii}}{s_i} \\
\langle \Phi | \hat{H}_{\mathrm{elec}} | \widetilde{\Phi} \rangle_i &= \sum_j \frac{\langle ij||ij \rangle}{s_j} + h_{ii} \\
\langle \Phi | \hat{H}_{\mathrm{elec}} | \widetilde{\Phi} \rangle_{ij} &= \langle ij||ij \rangle \\
\langle \Phi | \hat{H}_{\mathrm{elec}} | \widetilde{\Phi} \rangle_{ijk} &= 0
\end{aligned} \tag{2.34}$$

$$\mathbf{h} = \mathbf{C}\mathbf{h}^{\mathrm{HF}}\widetilde{\mathbf{C}}^{*} \tag{2.35}$$

Unlike in equation 2.26 where the matrix elements were classified by the number of excitations, here they are classified by the number of zero overlaps between

the biorthogonal orbitals, the notation $\langle\Phi|\hat{H}|\widetilde{\Phi}\rangle_i$ represents the matrix element of a determinant pair where the i^{th} biorthogonalized MOs of Ψ and $\widetilde{\Psi}$ have zero overlap. Zero overlaps will almost always be a result of electron excitations, hence the different forms of these equations can still be practically treated as representing different degrees of excitation between the determinants. The matrix $\mathbf{h}$ is the core Hamiltonian $\mathbf{h}^{\text{HF}}$ defined in equation 2.3 transformed into the biorthogonal-orbital basis

For computational efficiency we follow the Head-Gordon group, and implement the generalized Slater-Condon rules matrix form in terms of two different density matrices describing the common electron density between the two configurations $\mathbf{W}$ and $\mathbf{P}$.

Zero Overlaps	**Matrix Form**
None	$\frac{1}{2}(\mathbf{W}\cdot\mathbf{J}[\mathbf{W}]+\mathbf{W}\cdot\mathbf{K}[\mathbf{W}^{\alpha}]+\mathbf{W}\cdot\mathbf{K}[\mathbf{W}^{\beta}])+\Sigma_i\frac{\mathrm{h}_{ii}}{s_i}$
s_i	$\mathbf{P}_i\cdot\mathbf{J}[\mathbf{W}]+\mathbf{P}_i\cdot\mathbf{K}[\mathbf{W}^{\sigma_i}]+\mathrm{h}_{ii}$
s_i, s_j	$\mathbf{P}_i\cdot\mathbf{J}[\mathbf{P}_i]+\mathbf{P}_i\cdot\mathbf{K}[\mathbf{P}_i^{\sigma_i}]$
>2	0

Superscripts α and β denote quantities calculated using only the alpha or beta orbitals and σ_i denotes a quantity using only orbitals of the same spin as orbital i. The core Hamiltonian $\mathbf{h}$ is also transformed into the biorthogonal basis (equation 2.35). $\mathbf{W}$ is the unweighted density matrix and $\mathbf{P}_i$ the weighted density matrix constructed from the ith columns of the two $\mathbf{C}$ matrices.

$$\mathbf{P}_i = (\mathbf{C}_{\cdot i})^T\widetilde{\mathbf{C}_{\cdot i}} \tag{2.36}$$

$$\mathbf{W} = \sum_{s_i\neq 0}\frac{\mathbf{P}_i}{s_i} \tag{2.37}$$

Unlike orthogonal CI, in NOCI the nuclear part of the Hamiltonian doesn't necessarily become zero from the orthogonality of the MOs, which means we also need to add a term corresponding to the nuclear-repulsion energy scaled by the overlap of

the two sets of orbitals. Thus equation 2.38 is used to find the total energy of the Hamiltonian element.

$$\langle \Psi_i | \hat{H} | \Psi_j \rangle = S_{ij}^{\text{red}} H_{ij}^{\text{elec}} + S_{ij} E^{\text{nuc}} \tag{2.38}$$

Once the Hamiltonian matrix has calculated the NOCI coefficients $\mathbf{C}$ and the energies of the NOCI states $\mathbf{E}$ can be found by solving the generalized eigenvalue problem 2.39. Unlike in orthogonal CI, the overlap matrix $\mathbf{S}$ is unlikely to be the identity matrix.

$$\mathbf{HC} = \mathbf{ESC} \tag{2.39}$$

Natural Orbitals in SF-NOCI

We have found that the CUHF natural orbitals (equation 2.17) are a very effective set to use in SF-NOCI in place of the canonical HF orbitals, because they make the process of constructing and diagonalising the NOCI matrix more numerically stable.

However, the higher spin reference determinants generated by CUHF contain an active space of orbitals with occupation numbers of exactly 0.5 (singly occupied orbitals), which may not be Aufbau-ordered. Since - at HF level - all linear combinations of the singly occupied orbitals have equivalent energy, the precise character of these MOs is highly unstable to very small differences in the canonical MOs.

In SF-NOCI these orbitals are *not* equivalent; determinants with electrons in the orbital designated as the high-spin HOMO will be constructed using the high-spin MOs, while other determinants will use lower spin reference sets. Consequently the instability of the singlet orbitals can cause large discontinuities in the SF-NOCI potential energy surface.

We have found a solution to this is to use the linear combinations of the singly

occupied orbitals that overlap the least with the doubly occupied space of the lowest energy determinant. To find these we first define a matrix $\mathbf{B}$ of the overlaps between the singly occupied NOs $\mathbf{C}^S$ with the doubly occupied space $\mathbf{C}^D$.

$$\mathbf{B} = (\mathbf{C}^S)^T(\mathbf{S}\mathbf{C}^D)(\mathbf{S}\mathbf{C}^D)^T\mathbf{C}^S \tag{2.40}$$

Here $\mathbf{S}$ is the matrix of AO overlaps, $\mathbf{C}^S$ are the columns of the NO matrix with occupation numbers of 0.5 and $\mathbf{C}^D$ the columns with occupation numbers of 1. The eigenvectors of $\mathbf{B}$ can now be used to define a new set of singly occupied orbitals $\widetilde{\mathbf{C}}^S$ as linear combinations of the old singly occupied orbitals.

$$\mathbf{B}\mathbf{v} = \lambda\mathbf{v} \tag{2.41}$$

$$\widetilde{\mathbf{C}}^S = \mathbf{C}^S\mathbf{v}^T \tag{2.42}$$

We then sort the columns of $\widetilde{\mathbf{C}}^S$ in ascending order of their overlap with $\mathbf{C}^D$ and use them to replace the old $\mathbf{C}^S$ columns in the full NO matrix.

Chapter 3

SF-NOCI for Ground States

3.1 Introduction

In this chapter we will apply minimal-determinant SF-NOCI methods to two prototypical systems where static correlation effects dominate the shape of the potential energy curves: dissociating lithium hydride and twisting ethylene. These molecules have been chosen because they are relatively small and are common test cases for new quantum chemical methods, including those dealing with degeneracy or near degeneracy effects and molecular dissociation.[63,97] The advantage of using smaller molecules is that higher-level calculations are quick to perform; enabling rapid implementation and development, and benchmarking against alternative static correlation methods, such as MCSCF, using existing implementations available in standard quantum chemical program packages. Smaller molecules also allow the resulting MOs and wavefunctions to be interrogated in detail, yielding physically interpretable models of electron behaviour.

In both cases our goal is to use SF-NOCI to construct very simple and physically meaningful wavefunctions using a minimal set of determinants. To this end we are not attempting to completely recover the electron correlation energy, but rather to find a statically-correlated wavefunction that provides a useful starting point for

further dynamic correlation corrections.

3.2 Lithium Hydride

3.2.1 Model Requirements

LiH has been used as a test case for a wide range of new computational methods.[98–100] Because it has only 4 electrons, it can be easily benchmarked against high-level *ab initio* calculations and experimental data,[101] using large enough atomic orbital basis sets to eliminate inflexibility of the basis set as a potential error source. Furthermore, because the Li *s* orbital is far lower in energy than any other atomic or molecular orbitals in the system, these electrons play very little role in bonding. Therefore, LiH can be analysed as a pseudo-2-electron system, as differences between wavefunctions depend almost exclusively on the behaviour of the valence electrons. However, despite this simplicity, the electronic structure changes significantly over the dissociation curve. The ground state is ionic at bonding distances but dissociates into two neutral atoms.[102] Further, the ground and first excited singlet states experience an avoided crossing as the molecule dissociates.[102] Potential energy curves derived from spectroscopic measurements are illustrated in Figure 3.1.

These factors mean that many electronic structure methods struggle to represent the entire dissociation curve in a qualitatively correct manner.[103] In general, modelling dissociation requires a smooth transition from representing valence electrons using delocalized, doubly-occupied bonding orbitals to localising the bonding electrons on the dissociated fragments, in an appropriately spin-symmetric manner.

Ground state dissociation of LiH is also a particularly challenging case for single-reference SFCI as the fewer electrons present, the greater the relative difference between the spin flipped state the other electronic states. Since LiH has only 4 electrons the singly spin flipped state differs from the ground state in the spin and orbital occupancy of 25% of the electrons. Consequently, the singly spin flipped

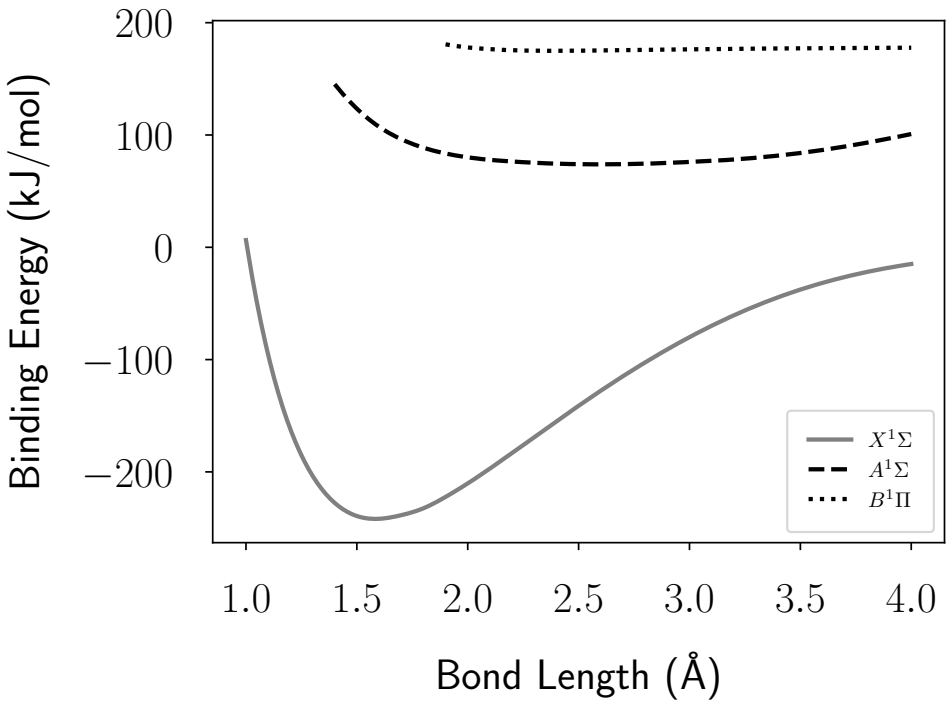

Figure 3.1: Experimental dissociation curves for the first 3 singlet states of LiH[101]

MOs are a worse basis to represent the ground state determinant than they would be in a larger system where a smaller proportion of the electrons are affected. This is a particular problem at equilibrium, where the singlet ground-state HF reference is a good first order approximation to the wavefunction, and the SFCI reference constructed from triplet-state MOs is particularly poor.

As discussed in the previous two chapters, the key to modelling the wavefunction with a minimal number of determinants is ensuring that the MOs making up those determinants are themselves physically meaningful, and as well-optimized for describing each electronic configuration as possible. In the case of dissociation, this means that the MOs have to be physically meaningful over the whole of the dissociation curve. The two important regimes are equilibrium (near bonding distance), where the wavefunction is dominated by the ground state HF singlet reference, and dissociation where the wavefunction will be fundamentally multi-configurational and reduce to appropriately antisymmetrized linear combinations of atomic orbitals. The intermediate regime will be a mixture of these two situations.

3.2.2 Computational Details

We performed calculations on LiH at bond lengths ranging from 1.0 Å to 4.0 Å, separated by 0.1 Å increments. Calculations were performed out using UHF, CASSCF(2,2), CCSD(T) SFS-CISD, and FR-SFS-NOCIS theories in conjunction with both STO-3G[104–106] and cc-pVTZ[107,108] atomic orbital basis sets. In addition, custom MCSCF calculations were carried out in a determinant basis including only the ground state configuration and the two spin-equivalent singly-excited states. STO-3G calculations were performed in order to generate relatively simple wavefunction expansions that can be easily illustrated and analysed, whereas the cc-pVTZ basis was used to generate accurate potential energy curves.

UHF, CCSD(T), and SFS-CISD calculations were carried out using the Q-Chem quantum chemistry package.[109] CASSCF(2,2) and MCSCF calculations were performed using the GAMESS package.[110] FR-SFS-NOCIS calculations were carried out using our in house "pychem" software developed as part of this thesis.

3.2.3 Reference Orbitals

Before discussing the nature of the multi-configuration wavefunctions, it is useful to examine the molecular orbitals that underpin the contributing determinants. As anticipated, the lowest energy MO in LiH is a doubly-occupied Li 1s orbital for all methods and geometries which is approximately 5300 kJ/mol lower in energy than the next lowest MO. From equilibrium to dissociation the energy of this core orbital changes by only 105 kJ/mol, compared to a 525 kJ/mol change in the energy of the HOMO. Consequently, this analysis will focus exclusively on the behaviour of the valence molecular orbitals.

Spin-restricted HF (RHF singlet, CUHF triplet) and spin-unrestricted UHF dissociation curves are shown in Figure 3.2, while Figure 3.3 illustrates RHF and CUHF frontier MOs obtained at equilibrium and at dissociation.

At equilibrium, the RHF HOMO is a σ bonding orbital dominated by the in-phase

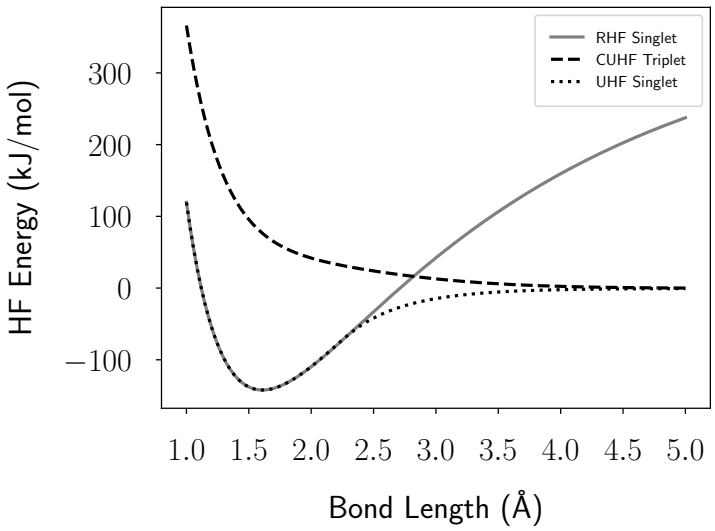

Figure 3.2: The HF energy of the ground state single and triplet states of LiH in the cc-pVTZ basis. The singlet is calculated using the RHF and UHF methods and the triplet only with CUHF.

mixture of the H 1s, Li 2s and Li p_z AOs. At dissociation, this bonding orbital retains its delocalized character. This results in the "classic" RHF behaviour of the potential energy curve. The dissociation energy is significantly overestimated, due to the wavefunction comprising of a mix of covalent/atomic and ionic terms. This is also reflected in the Mulliken atomic charges of $\pm 0.495e$.

In contrast, the CUHF triplet orbitals are much more atomic in character over the whole of the dissociation curve. Although not immediately apparent from Figure 3.3, both singly-occupied CUHF orbitals are primarily atomic in character even at equilibrium. The ROHF HOMO is dominated by the LiH 2s AO with a coefficient of 0.93 and HOMO-1 by the H 1s AO with a coefficient of 0.85. The remaining AO contributions to each MO give these MOs the observed bonding and antibonding character. The net effect of these MOs being more "atomic-like" than "covalent-like" is that the triplet state is unbound, *i.e.* much higher in energy than the singlet at equilibrium, and higher in energy than the sum of the energies of the two isolated atoms. These molecular orbitals become even more atomic in character as the molecule dissociates. The coefficients of dominant AOs reach 1.0 by 4.0 Å, indicating localisation of same-spin electrons to different atomic centres.

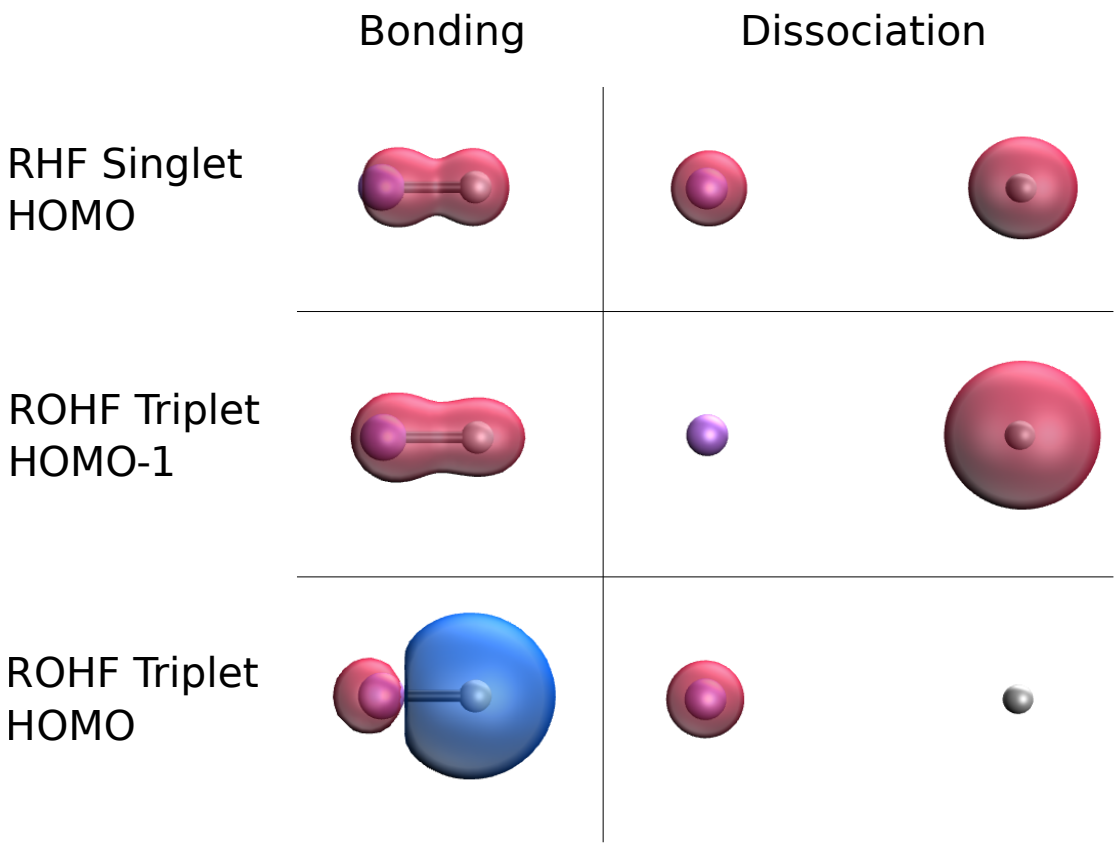

Figure 3.3: The frontier MOs of LiH at both bonding (1.6 Å) and dissociation (4.0 Å) All MOs calculated using restricted HF in the cc-pVTZ basis set

The UHF wavefunction resembles the RHF wavefunction at equilibrium but the CUHF wavefunction at dissociation, because at infinite separation there can be no Fermi repulsion, and so spin-labelling is unimportant when it comes to optimizing the shape of the orbitals. Therefore, the UHF curve in Figure 3.2 smoothly transitions from RHF-like at equilibrium to the correct dissociation limit. However, this comes at the cost of breaking spin symmetry. While it's appropriate for opposite-spin electrons to localize onto separate atomic centres, forming a single spin-symmetric determinant, it's not appropriate for same-spin electrons, which should be indistinguishable but become distinguishable upon localization. This is known as the "spin symmetry breaking problem" in UHF, generally prohibiting UHF being used as a foundation for more rigorous methods that account for static correlation *and* obey spin symmetry.

Between the RHF singlet and the CUHF triplet molecular orbital sets we can form

reference determinants that appropriately model the wavefunction at equilibrium or at dissociation, but no single-reference that can do both. Therefore, both sets are combined in the FR-SFS-NOCIS model to form a determinant basis that should be - in principle - capable of recovering the RHF wavefunction at equilibrium, and the antisymmetrized spin-flipped CUHF wavefunction at dissociation, and smoothly interpolating between them, maintaining the correct spin symmetry throughout.

3.2.4 Potential Energy Curves and Multi-Reference Wavefunction Analysis

The performance of the FR-SFS-NOCIS model for modelling LiH dissociation in the ground state is benchmarked against UHF, SFS-CISD, MCSCF, CASSCF(2,2), CCSD(T) and experiment in Figure 3.4. The experimental curves were constructed from vibrational spectroscopy data using the inverted perturbation approach.[101,111]

From Figure 3.4, it is evident that only the FR-SFS-NOCIS, MCSCF, CASSCF(2,2) and CCSD(T) models produce qualitatively correct dissociation curves. CCSD(T)/cc-pVTZ almost exactly reproduces the experimental curve, due to its ability to concurrently recover both static and dynamic contributions to the total correlation energy.

The advantages and deficiencies of the UHF model are well known. Its primary advantage lies in its ability to smoothly transition between the RHF wavefunction and energy at equilibrium and yield the dissociation energy as the sum of the energies of the isolated fragments. However, this comes at the expense of breaking spin-symmetry, which results in the UHF dissociation curve taking on a qualitatively different shape to the others, particularly in the "intermediate" bond length regime ($R \sim 2R_e$).

To better understand the multi-reference models, it is instructive to compare wavefunction composition as a function of bond length. CI coefficients for the FR-SFS-NOCIS, SFS-CISD, MCSCF and CASSCF(2,2) models at equilibrium ($R = 1.6$ Å),

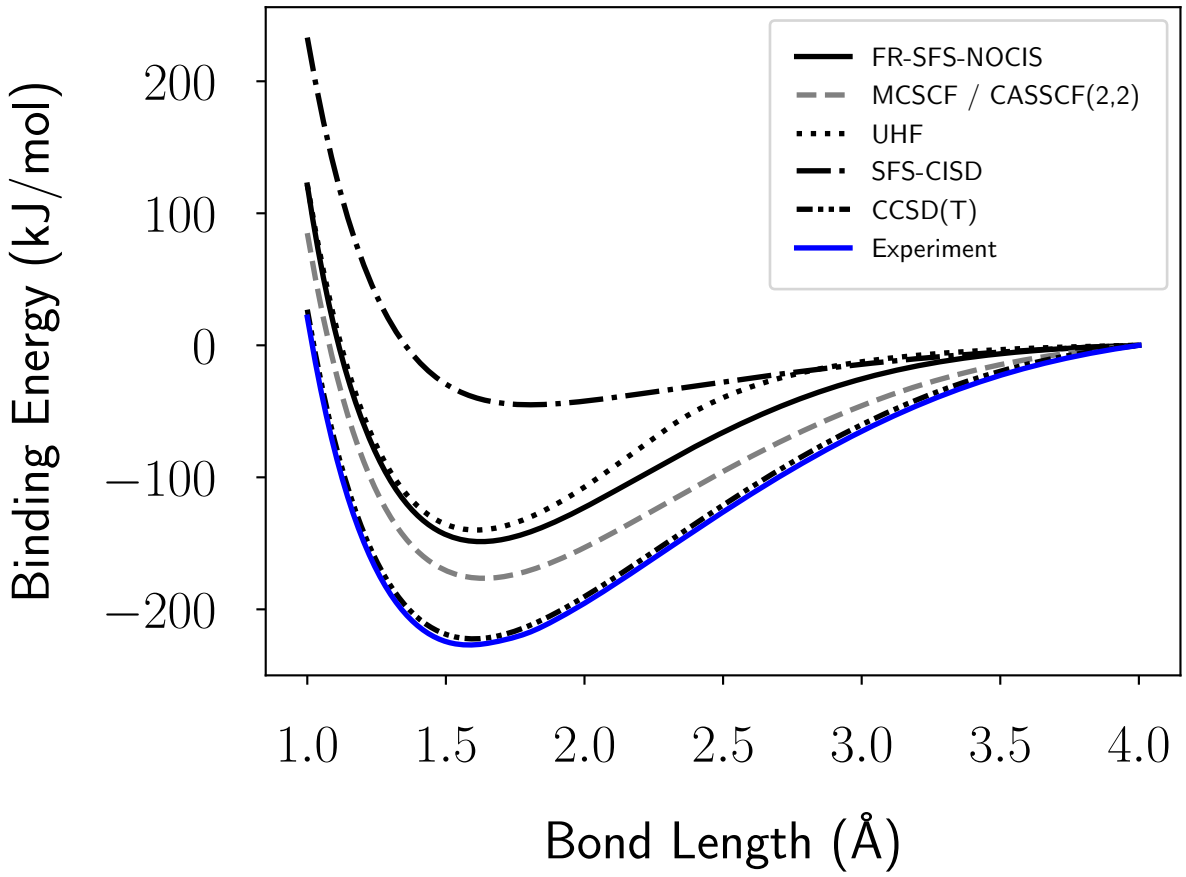

Figure 3.4: Dissociation curves of ground state LiH calculated at five different levels of theory, in the cc-pVTZ basis set, and the spectroscopically-determined experimental curve.[101] CASSCF and MCSCF energies are the same to within the line width resolution. All curves are shifted up by E(H) + E(Li) to ensure that they tend to the same dissociation limit.

Method	Determinants	Ref	Equilibrium	Intermediate	Dissociation
FR-SF-NOCI	$1\sigma^2 2\sigma^2$	S	0.9455	0.6008	0.0000
	$1\sigma^2 2\sigma^1 3\sigma^1$	T	0.1199	0.4202	0.7072
SFS-CISD	$1\sigma^2 2\sigma^2$	T	0.7765	0.5195	0.0000
	$1\sigma^2 2\sigma^1 3\sigma^1$	T	-0.4422	-0.5918	0.7059
	$1\sigma^2 3\sigma^2$	T	0.0000	-0.1723	0.0000
CASSCF	$1\sigma^2 2\sigma^2$	$M_{(2,2)}$	0.9898	0.8490	0.0098
	$1\sigma^2 2\sigma^1 3\sigma^1$	$M_{(2,2)}$	-0.1416	0.2133	0.7021
	$1\sigma^2 3\sigma^2$	$M_{(2,2)}$	-0.0111	-0.4337	-0.0667
MCSCF	$1\sigma^2 2\sigma^2$	$M_{(1,2)}$	0.8482	0.5859	0.0316
	$1\sigma^2 2\sigma^1 3\sigma^1$	$M_{(1,2)}$	-0.3746	-0.5730	-0.7078

Table 3.1: The CI coefficients for each of the multi-configuration methods at three different points along the dissociation curve: equilibrium (1.6 Å), intermediate (3.0 Å) and dissociation (4.0 Å). In each open shell states the two spin-symmetric configurations were in phase. The **Ref** column describes the reference configuration used to generate the MOs used in that determinant, S denotes the RHF singlet, T the CUHF triplet, $M_{(2,2)}$ the (2,2) active space MCSCF and $M_{(2,2)}$ the MCSCF containing the ground configuration plus two HOMO to LUMO single excitations. All configurations not shown are 0

stretched ($R = 3.0$ Å) and dissociated ($R = 4.0$ Å) bond lengths are listed in Table 3.1. For the new FR-SFS-NOCIS method, the CI coefficients are also plotted as a function of bond length in Figure 3.5

Equilibrium

The equilibrium behaviour of the FR-SFS-NOCIS and CASSCF(2,2) models contrast sharply with that of SFS-CISD, both in terms of the predicted binding energies and wavefunction composition. Both FR-SFS-NOCIS and CASSCF(2,2) wavefunctions are dominated by contributions from a determinant with a doubly-occupied bonding orbital. However, the SFS-CISD wavefunction contains substantial contributions from singly-excited states even at equilibrium, and dramatically underestimates the binding energy.

Overall, this indicates that the SFS-CISD wavefunction is including these contributions in an attempt to ameliorate the poor quality of the triplet-reference MO

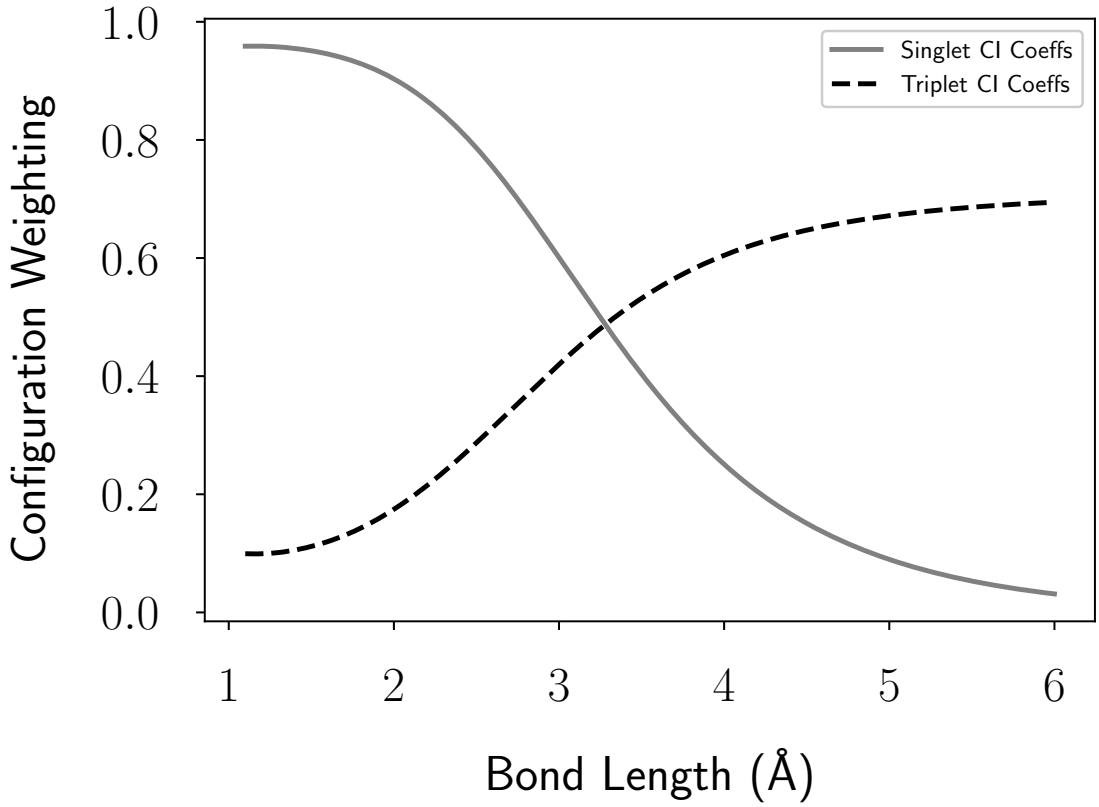

Figure 3.5: The change in CI coefficients over the dissociation curve of LiH for the two determinants used in a RF-SFS-NOCIS calculation cc-pVTZ basis set. The singlet determinant begins dominant at bonding distances and its contribution smoothly decreases as the bond breaks. Note the "Triplet CI Coeffs" curve represents the coefficients of two determinants in the CI wavefunction - the two spin-symmetric singly-excited determinants.

basis for expanding a state that is much better represented by a simple singlet-reference determinant, *i.e.* starting from a basis that includes two electrons in a single well-optimized bonding orbital.

This illustrates a key limitation of the SFS-CISD model - it cannot be applied as a minimal-determinant model when the reference set of high-spin MOs does not closely resemble the corresponding low-spin orbital set for the MOs that are commonly occupied across both spin states. The conventional solution to this problem is simply to include additional determinants, constructed by allowing excitations into the virtual orbital space. However, in moving from a minimal-determinant model to a more extensive CI model, the physical interpretability of the wavefunction is lost.

The FR-SFS-NOCIS approach clearly ameliorates this deficiency. Using different spin-reference determinants and selecting the most appropriate MO set for each configuration allows us to include the ground-state Hartree-Fock configuration with fully optimized doubly-occupied bonding orbitals in addition to singly-excited states formed from the triplet reference MO set. Clearly, this is a more appropriate determinant basis for expanding the wavefunction at equilibrium, where the conventional HF model is already a good 0th order approximation to the true wavefunction. This assertion is supported by the fact that the FR-SFS-NOCIS and CASSCF(2,2) wavefunctions are both dominated by contributions from this determinant at equilibrium, and also that the UHF, FR-SFS-NOCIS and CASSCF(2,2) models are in broad qualitative agreement about the magnitude of the binding energy.

Closer inspection of the FR-SFS-NOCI and CASSCF(2,2) potential energy curves and CI wavefunction coefficients reveals a somewhat surprising feature - the CASSCF(2,2) wavefunction is even more strongly dominated by the singlet HF reference. This is surprising because the FR-SFS-NOCI wavefunction contains the HF singlet determinant, which is optimized to be the best possible single-configuration description of the wavefunction, while the CASSCF(2,2) MOs are optimized for a multi-configuration description of the wavefunction. Therefore, we would expect the CASSCF wavefunction to be more multi-configurational than the FR-SFS-NOCI. One potential explanation for this is that the inclusion of the doubly-excited determinant allows the CASSCF(2,2) model more flexibility and therefore enables it to recover more of the binding energy. However, this configuration has a small coefficient, which suggests that this might not be the entire explanation.

The other explanation is, that orbital optimization may allow the singly-excited states to recover more of the correlation energy by concurrently optimizing the shape of the original LUMO while also optimizing how these states mix in to the overall wavefunction. This supposition is supported by a number of lines of evidence:

1. The singly excited state CI coefficients get larger and change sign moving

from FR-SFS-NOCIS to CASSCF(2,2), changing from an in-phase combination with the ground state reference determinant to out-of-phase.

2. MCSCF results continue this trend. Although the MCSCF wavefunction composition is very different to the CASSCF(2,2) wavefunction composition, the energies are the same. This indicates that additional correlation energy can be recovered by either including the doubly-excited determinant or optimizing the singly-excited determinants, and that these two quite different approaches have the same energetic effect in this particular case.

3. Even for the FR-SFS-NOCIS model in which the excited state determinant MO set is only partially optimized, this determinant contributes substantially to the wavefunction and lowers the binding energy substantially relative to the single-determinant UHF model.

Overall, it is most likely that these singly-excited state configurations are capturing intra-atomic correlation effects on the Li atom. However, more work is required to get to the bottom of these observations, and understand their physical origin. In particular, the key question becomes "are these truly minimal-determinant models or are the singly-excited determinants capturing some dynamic correlation effects?". However, for practical purposes, this is not particularly important, because all that we need is reliable black-box static correlation model that does not scale as badly with active space size as CASSCF and does not require the user to carefully tailor the active space and monitor its composition like MCSCF.

At equilibrium, this means that any proposed model must yield a wavefunction that is dominated by the HF reference determinant, for single-reference systems. FR-SFS-NOCIS clearly fits this bill.

Intermediate

In the intermediate regime, the CASSCF, MCSCF and FR-SFS-NOCI models all have qualitatively similar behaviour energetically despite having quite different sets

of optimized (or semi-optimized) MOs and CI coefficients. Interestingly, only the FR-SFS-NOCIS and MCSCF models have CI coefficients that change monotonically with bond length.

Once again, this implies redundancy within the CASSCF(2,2) active space, with both the singly- and doubly- excited configurations contributing to recovering the same "chunk" of correlation energy. In the intermediate regime, this has a plausible physical interpretation - it appears that the dissociating bond is being modelled using a combination of molecular orbital (MO) and valence bond (VB) theories. The doubly-occupied bonding and antibonding orbitals describe bond dissociation according to MO theory, while the singly-occupied atomic-like orbitals form VB-like contributions to the wavefunction. Because of this redundancy, it appears that there are many iso-energetic ways of mixing these two sets of contributions (plus an additional contribution from the doubly-occupied bonding term to account for the fact that in the intermediate regime the molecule is still partly bound / not fully dissociated), including dropping the doubly-excited configuration entirely, yielding the MCSCF model.

It is instructive to compare the MCSCF and FR-SF-NOCIS models, because they use the same determinant basis and differ only in whether the underlying orbitals are fully optimized and orthogonal (MCSCF) or partially optimized via different spin-reference HF calculations and non-orthogonal (FR-SF-NOCIS). Again, it is clear that excited state orbital optimization can make quite a significant difference to the amount of correlation energy that can be recovered.

Nonetheless, FR-SFS-NOCIS remains an appropriate static correlation model - well and uniquely defined, with a smooth potential energy curve that is largely parallel to the more complete MCSCF and CASSCF curves, and with CI coefficients that also smoothly and monotonically change as a function of bond length.

Dissociation

For LiH, all multi-reference models considered here become equivalent at dissociation, formed as an antisymmetrized linear combination of equally-weighted singly-excited reference states formed from localized atomic orbitals. For CASSCF in particular, this is different to what is observed for H_2, the canonical exemplar, in which the wavefunction is described as a linear combination of states formed from doubly-occupied bonding and antibonding orbitals.

Overall, both the FR-SFS-NOCIS and MCSCF models can be described as models that smoothly interpolate from a delocalized MO-dominated description of the wavefunction at equilibrium to a localized AO valence-bond-like wavefunction at dissociation. MCSCF recovers more correlation energy at equilibrium and along the dissociation curve by using fully optimized orbitals for singly-excited determinants, but this comes at the expense of needing to concurrently optimize both MO and CI coefficients.

The greatest advantage of SF-NOCI models compared to "conventional" multi-configurational SCF methods is their black-box nature. They are not sensitive to the initial choice of active space or orbital ordering, and do not require user input or careful validation.

However, away from the dissociation limit, there is a fundamental mismatch how well the high-spin and low-spin reference orbitals describe the corresponding NOCI states. Clearly, the singlet ground state will be better described by singlet reference orbitals, whereas the singlet excited states within the NOCI expansion rely on triplet HF reference orbitals. Further, the decoupling of the orbital and CI coefficient optimization processes means that neither set of coefficients is fully optimized.

3.2.5 Correlation Energy Analysis

Overall, a good static correlation model will provide an appropriate starting point for recovering the remaining dynamic correlation energy. For a simple system like

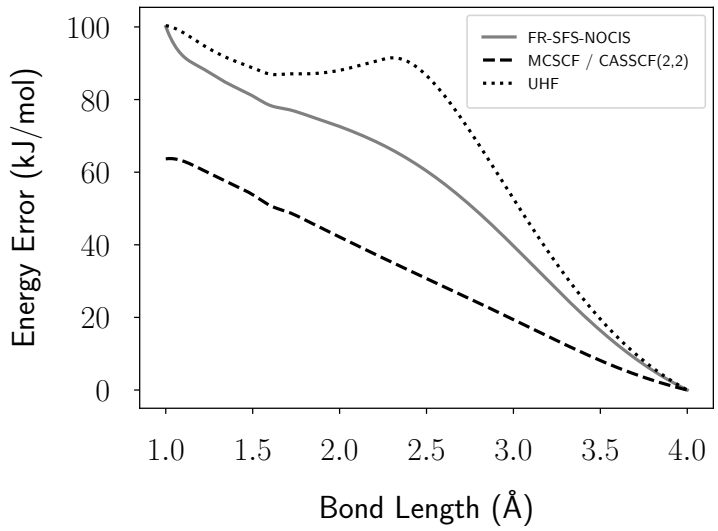

Figure 3.6: The deviation in PP-SFS-NOCI, UHF and MCSCF energy from experiment over the course of the dissociation curve of LiH.

LiH with only two valence electrons, the dynamic correlation energy is expected to decrease monotonically with bond length, as the two electrons get further apart.

To test the utility of FR-SFS-NOCIS, MCSCF and UHF as static correlation models, the remaining dynamic correlation contribution that would need to be recovered to match experiment is illustrated in Figure 3.6.

Clearly, the MCSCF correlation energy curve displays the expected behaviour - the correlation energy decreases smoothly and monotonically from equilibrium to dissociation.

The UHF curve differs significantly from the MCSCF curve in the intermediate regime. Here, the MCSCF wavefunction is most multi-configurational in character, with all three underlying determinants contributing approximately equally to the wavefunction at 3.3 Å. This implies that the rise in the UHF difference curve is caused by UHF being unable to appropriately account for the multi-configurational nature of the wavefunction in this regime.

SFS-NOCIS has a similar but much smaller rise in the same region of the curve. This likely indicates that the spin flipped MOs are not ideal in this regime. At equilibrium the singlet MOs describe the wavefunction very well and at dissociation the

triplet MOs are very good. In the intermediate range using both sets is better than using either one individually optimized but is still not as good as the concurrently optimized MOs used by MCSCF based methods.

3.3 Ethylene

3.3.1 Model Requirements

Twisting ethylene is commonly used as a test case for multi-reference methods.[60,97] As the C-C double bond twists the π and π^* MOs become degenerate at 90°, breaking the π bond and meaning the single π^2 determinant cannot appropriately model the wavefunction - as it can in the planar conformation.[112] The resulting system is considered one of the prototypical cases of a diradical system,[113] with two electrons previously involved in the π bond each localized onto one of the carbon atoms.

In the following discussion, the planar ethylene molecule is assumed to lie in the xy plane, with the C-C bond along the x axis. The angle θ refers to the dihedral angle between two cis hydrogen atoms as shown in Figure 3.7, *i.e.* $\theta = 0°$ in the planar conformation.

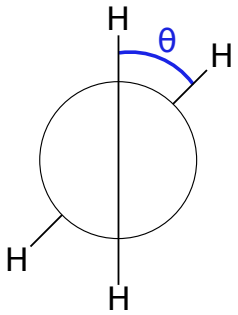

Figure 3.7: A Newman projection looking down the C-C bond of ethylene (i.e. along the x axis), showing the definition the angle θ used in the text

As the bonding and anti-bonding orbitals become degenerate most methods fail to effectively model the region around the top of the rotation barrier. An effective minimal determinant model would require only the π^2 and π^{*2} configurations, with both of these determinants contributing equally to the wavefunction at 90°, effectively

localizing the two valence electrons onto the two carbon atoms to form a diradical species.

The experimentally derived torsion energy for ethylene is approximately 272 kJ/-mol,[114] so cannot be compared directly to computed barrier heights that do not account for zero-point and/or thermal vibrational effects.[115,116] Hence, we do not expect to exactly replicate the experimentally-derived torsional barrier height and will use a CCSD(T) potential energy curve to benchmark against.

3.3.2 Computational Methods

The equilibrium geometry of ethylene was optimized at CCSD(T)/cc-pVTZ using Q-Chem. The electronic energy was then calculated as the C-C bond was rotated in 5° increments from 0° to 85°, where 0° has all the atoms in a single plane and 90° generates a perpendicular conformation with both ends of the molecule geometrically orthogonal. Calculations were not performed at 90° due to convergence problems, to help account for this an additional calculation was performed at 88°. All bond lengths were held constant as the C-C bond was twisted.

These calculations were carried out using the SF-NOCI, SF-CISD, CASSCF, CCSD(T) and RHF methods. The SF-NOCI calculations where carried out using a basis of two determinants; the lowest energy configuration and the HOMO to LUMO doubly-excited configuration - making this a PP-SFS-NOCI calculation, or equivalently SFS-NOCID. CASSCF and SFS-CISD calculations were carried out using four determinants - the two included in PP-SFS-NOCI plus two spin symmetric single excitations. However, these configurations do not mix with doubly-excited determinants, by symmetry.

As with LiH the calculations were carried out in both the STO-3G basis set - for analysing the wavefunction - and the cc-pVTZ basis set - for comparing the energies.

3.3.3 Reference Orbitals

The HF energies of the singlet and triplet states are shown in Figure 3.8. As shown above, RHF overestimates the barrier for the singlet state, with a cusp at 90°.[66] The CUHF triplet curve has very different behaviour, reaching an energy minimum at 90°, more stable than the RHF singlet.

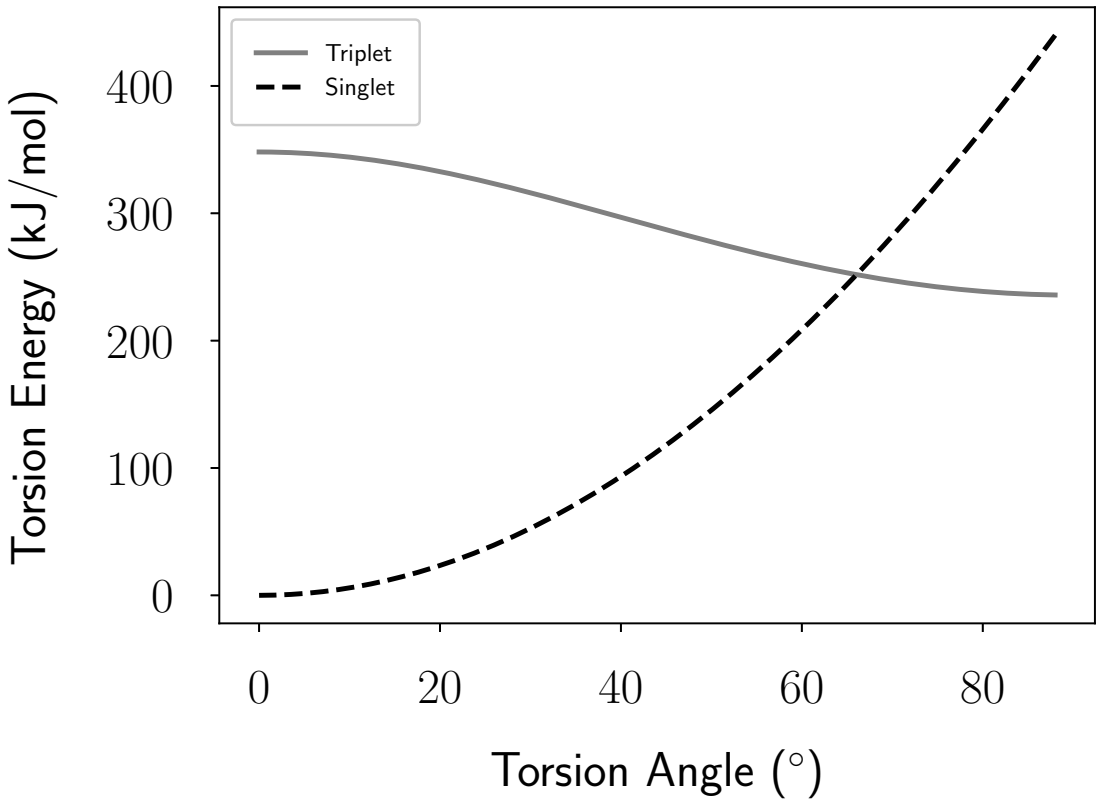

Figure 3.8: The HF energies of the RHF singlet and CUHF triplet ground states of of ethylene, calculated in the cc-pVTZ basis set. Energies are scaled so the singlet energy is zero in the planar conformation

Figure 3.9 shows the frontier orbitals of ethylene's singlet and triplet states in both configurations. At 0° the singlet HOMO is - as expected - a π bonding orbital. The carbon p_z orbitals each have coefficients of 0.63 and all other AOs have negligible coefficients (less than 10^{-14}). By 88° the dominant orbitals are an in phase combination of the p_z orbital on one carbon and the p_y orbital on the other carbon - each with coefficients of 0.67. Both these p orbitals are orthogonal to the plane formed by the carbon they are located on and its bonded hydrogens. Significant electron density has moved to the hydrogens which now have coefficients of ± 0.2.

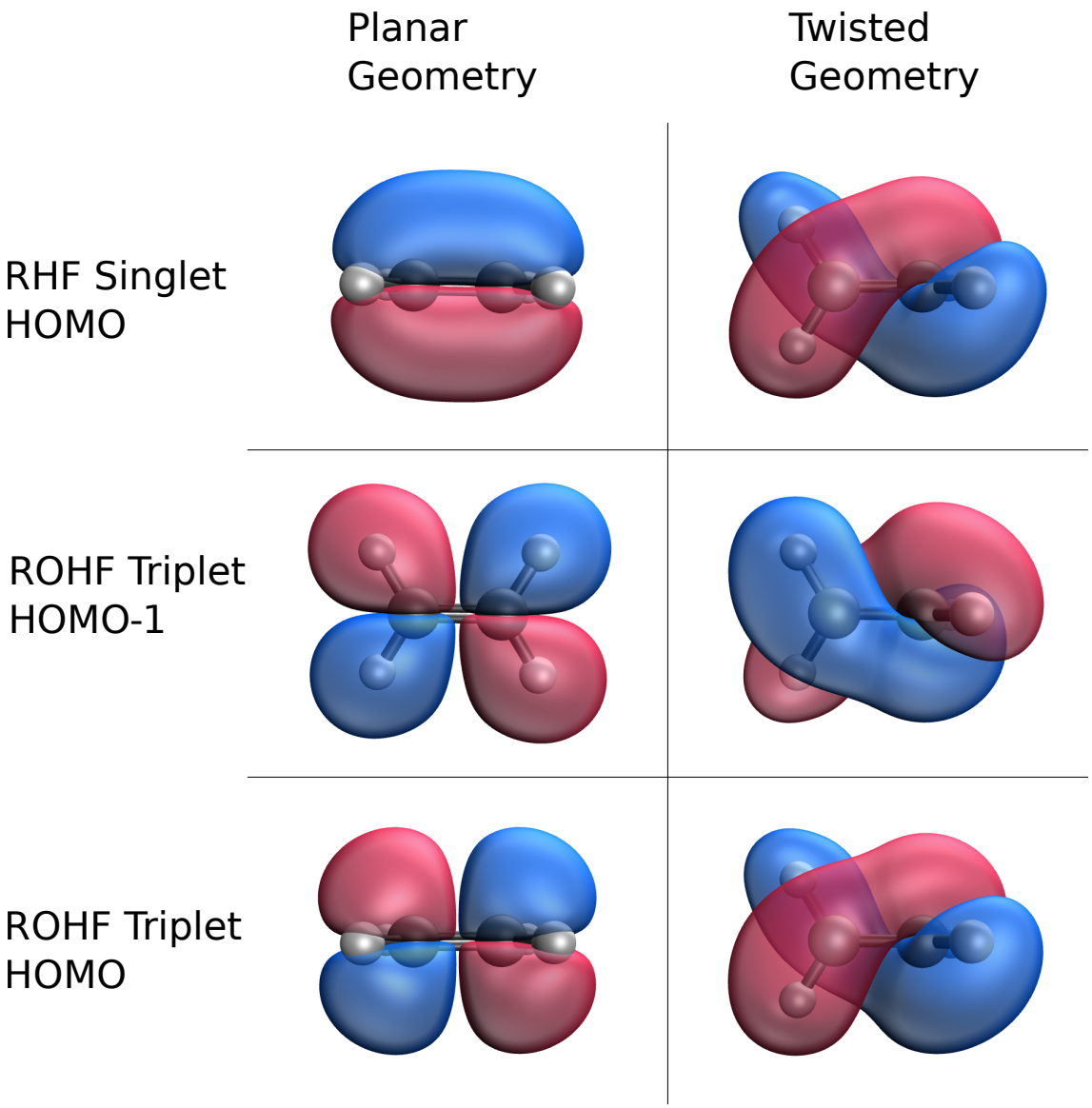

Figure 3.9: The frontier MOs of ethylene calculated in both the planar and twisted geometries in restricted HF and the cc-pVTZ basis set. Note the different orientation of the molecule in the ROHF triplet HOMO case to better display that MO.

The electron density moving into the two orthogonal p orbitals indicates the diradical nature of the system however - as with dissociation - RHF cannot effectively localize the electrons. Consequently the RHF wavefunction has some artefactual ionic character.

As shown in Figure 3.9, in the planar conformation the HOMO-1 is an π^* orbital comprising two out of phase p_z orbitals with coefficients of $\pm$ 0.81 and zero coefficients for all other AOs. The HOMO is dominated by the hydrogen $1s$ orbitals - which all have coefficients of $\pm$0.34 - and their in phase interaction with the car-

bon p_y orbitals - which have coefficients of ± 0.39. In the twisted conformation, the triplet MOs closely resemble the singlet MOs at the same geometry. In each case, the p orbital orthogonal to the bonded hydrogen atoms has the largest coefficient (≈ 0.66). However in the HOMO-1, these two orthogonal p orbitals are in phase with each other while in the HOMO the AOs are out of phase.

3.3.4 Potential Energy Curves and Multi-Configuration Wavefunction Analysis

Figure 3.10 shows the potential energy surfaces computed using RHF, CASSCF, CCSD(T), SFCI and PP-SFS-NOCI methods. All five methods have very similar energies up to about 50°, when the failure of RHF to describe the multi-configurational nature of the wavefunction starts to become more apparent. The four remaining methods are very similar until about 75°, when SFS-CISD begins to underestimate the barrier height relative to CASSCF and PP-SFS-NOCI, which overestimate the height relative to CCSD(T).

All four of the CI expansion based methods largely avoid the cusp at 90° predicted by RHF, although a slight cusp is apparent in the CCSD(T) curve.

The differences between the methods are more clearly seen when plotted as their deviation from the CCSD(T) energy - as in Figure 3.11. PP-SFS-NOCI slightly underestimates the torsion energy compared to CASSCF for most angles, however at 88° the two curves differ by only 2.3 kJ/mol, and underestimate the torsion energy by approximately 45 kJ/mol. SFS-CISD is similar to CASSCF for most of curve, but diverges sharply near 88° resulting in a 75 kJ/mol underestimation of the torsion energy - almost the same amount RHF overestimate the energy by (76 kJ/mol).

The fact that PP-SFS-NOCI predicts a very similar barrier height to CASSCF, while RHF and SFS-CISD diverge from the CASSCF curve close to 90°, suggests that multi-reference spin flipped orbitals provide a suitable approximate replacement for their multi-reference optimized counterparts, and therefore yield very similar CI

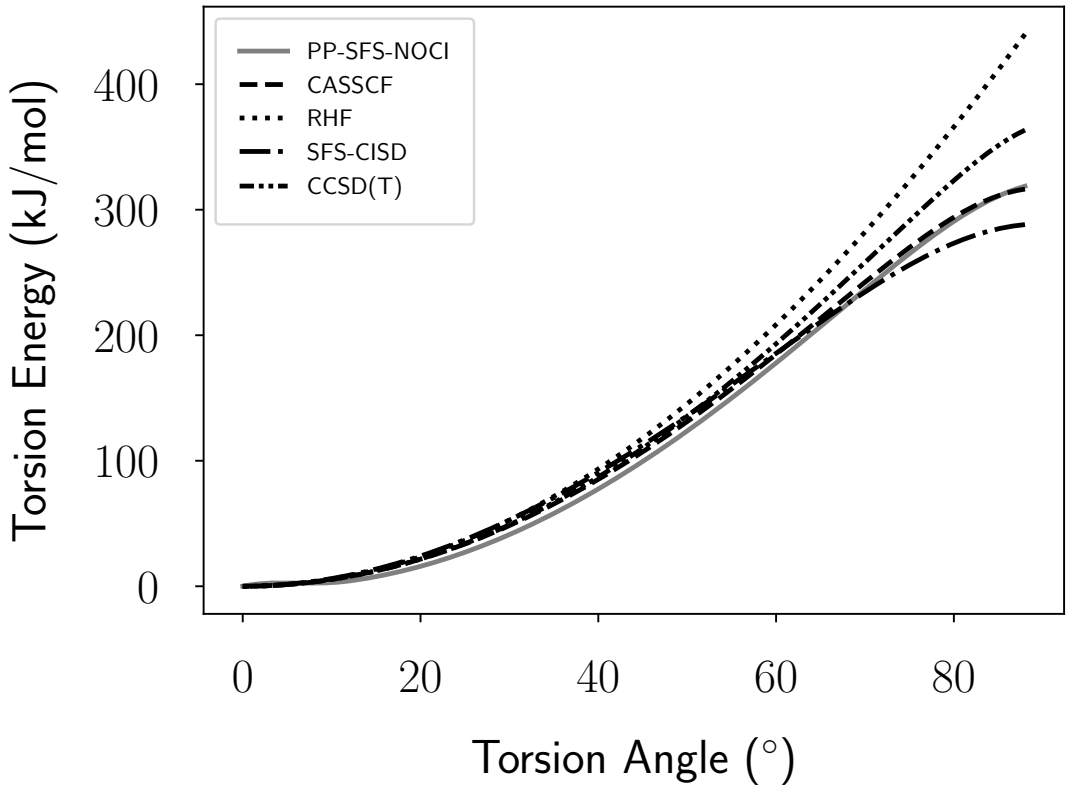

Figure 3.10: The potential energy surfaces for ethylene torsion calculated using the RHF, CASSCF, SSF-CISD, CCSD(T) and PP-SFS-NOCI methods in the cc-pVTZ basis set. All methods have been shifted so that their energy is zero in the planar conformation.

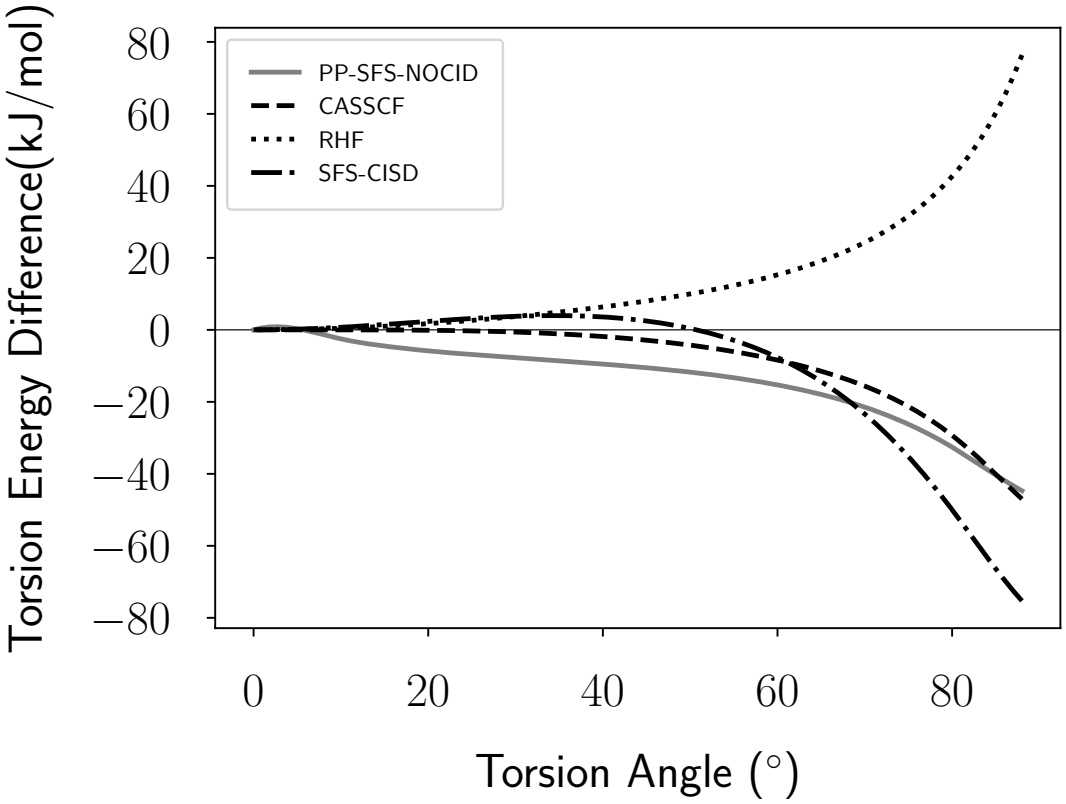

Figure 3.11: The difference in torsion energy of Ethylene for RHF, SFS-CISD, CASSCF and PP-SFS-NOCI compared to CCSD(T), calculated in the cc-pVTZ basis set.

Method	Determinant	Ref	Planar	Intermediate	Twisted
PP-SFS-NOCI	$(\pi)^2$	S	0.9792	0.9684	0.7933
	$(\pi^*)^2$	T	-0.2030	-0.2493	-0.6088
SF-CISD	$(\pi)^2$	T	0.9752	0.9631	0.7339
	$(\pi^*)^2$	T	-0.2215	-0.2692	-0.6792
CASSCF	$(\pi)^2$	$M_{(2,2)}$	0.9880	0.9767	0.7460
	$(\pi^*)^2$	$M_{(2,2)}$	-0.1542	-0.2146	-0.6659

Table 3.2: CI coefficients for the ground state of ethylene, in the planar (torsion angle of 0°), intermediate (45°) and twisted (88°) geometries. The **Ref** column describes the reference configuration used to generate the MOs used in that determinant, S denotes the RHF singlet, T the CUHF triplet and $M_{(2,2)}$ the (2,2) active space MCSCF. All configurations not shown are 0

wavefunction expansions.

The ground state CI coefficients for the different configuration wavefunctions are shown in Table 3.2. In the planar configuration, PP-SFS-NOCI, SF-CISD and CASSCF all have practically identical CI coefficients. For all methods, the wavefunction is dominated by the $(\pi)^2$ configuration with a weighting of ≈ 0.98 and a small contribution of $\approx$ -0.2 from from the $(\pi^*)^2$ configuration.

At 45°, the CI coefficients are still qualitatively the same as at 0°. All three wavefunctions become more multi-reference in character, although the CASSCF coefficients change the most.

At 88° there is slightly more divergence between the CI coefficients of the three methods, although they all remain in broad agreement. Overall, this indicates that the spin-flipped MO sets closely resemble their CASSCF counterparts. This confirms that the spin flip approach is a better approximation in larger symmetric molecules - like ethylene - than smaller asymmetric molecules like LiH.

However, PP-SFS-NOCI remains a better model than SFS-CISD because it more closely reproduces the CASSCF torsional potential energy curve. Clearly, there is a significant advantage in retaining HF-optimized orbitals for the ground state reference determinant.

3.3.5 State Switching

In the CASSCF and SFS-CISD cases, at the 88° point, the lowest energy state was not the state composed of the ground state and doubly-excited configurations but rather the state made up of equal contributions from the two singly-excited configurations (the so called T excited state of ethylene).[117]

In this case, this state switching is obvious, but in larger calculations with more determinants involved it could be easily missed. This illustrates an advantage of the minimal determinant approach - using few determinants minimizes the chance of finding non-target states. SF-NOCI could also capture these triplet states by using a wavefunction which includes single excitations, this will be explored in Chapter 5.

3.4 Conclusions

For LiH and twisting ethylene, SF-NOCI is able to predict qualitatively correct ground state dissociation and torsional potential energy curves, respectively. SF-NOCI outperforms orthogonal SFCI and yields results comparable to MCSCF-based methods. This demonstrates that SF-NOCI is capable of overcoming the fundamental problem with SFCI - the poor quality of the spin flipped orbitals - to describe states dominated by lower energy configurations.

These examples also illustrate two weaknesses with the SF-NOCI approach. Firstly, it may not fully recover static correlation in cases where none of the spin flipped states provide an ideal basis for the wavefunction - as is shown in the intermediate region of LiH dissociation. Secondly, the use of multiple reference determinants may not correctly model relationship between determinants that are enforced by symmetry, such as in the case of ethylene, where the π and π^* orbitals do not become degenerate across the singlet and triplet determinants - causing the wavefunction to deviate slightly from its expected form.

However, despite these weaknesses, the non-orthogonal CI approach gives energies

and wavefunctions that are qualitatively very similar to the more rigorous but expensive MCSCF method. These wavefunctions have the expected energies and behaviours without any need to include unphysically motivated determinants in the CI expansion or to do expensive multi-reference optimizations. Importantly, SF-NOCI models are completely defined by spin-flip level and excitation-level, *i.e.* they are well-defined black-box models that do not require user input.

However, like all static correlation models, the very small CI expansions used in SF-NOCI fail to include significant amounts of dynamic correlation, necessitating additional "post-CI" corrections to accurately recover total correlation energies.

Chapter 4

Excited State Models

Until now, this thesis has focused exclusively on ground electronic states, but neither the single-determinant molecular orbital model nor multi-reference CI and CASSCF models are limited to modelling the lowest energy electronic state. Practicalities of numerical convergence notwithstanding, none of the assumptions underlying the derivation of HF equations are exclusive to ground states and a number of attempts have been made to apply HF theory to higher-energy states.[31,65] Models that can accurately solve the Schrödinger equation for excited states are a very active area of quantum chemical research.[118–120] In practice, most excited state methods simply involve taking the higher-energy solutions to "conventional" methods like CI and CASSCF and treating those as representations of the excited states.

4.1 Existing Approaches

4.1.1 Configuration Interaction Singles

Since the CI equations yield multiple solutions to the time-independent Schrödinger equation, higher-energy eigenvectors can be interpreted as representing the excited states of a system. This allows CI theory to be trivially applied to model excited states, up to the same number of excited states as there are determinants in the CI

expansion.

The simplest such model is configuration interaction singles (CIS), in which a CI calculation is performed using all single excitations from a HF ground state reference. CIS has the advantage of being straightforward and computationally cheap; but fails to account for either static or dynamic electron correlation, which can lead to large errors in excitation energies, especially for transitions to or from near-degenerate states.

Further, since the reference orbitals are optimized for the ground state electronic configuration, CIS inherently provides a better quality description of the ground state than any of the excited states.[121] This is a problem with all single-reference excited state methods, but the limited number of determinants included in CIS makes this effect more pronounced, as double and higher excitations are not available to correct for the poor quality of the unoptimized virtual MOs used in constructing the excited state determinants.

This unbalanced description leads CIS to routinely overestimate excitation energies by 100-200 kJ/mol,[122] and even more in cases like charge transfer states, where the ground state MOs provide a particularly inappropriate basis.[123]

Because the CIS model does not recover any form of correlation energy, it also often fails to predict the correct ordering of excited states.[124] This is largely due to the differential importance of the neglected correlation contributions. For example, static correlation plays a large role in determining the relative energies of near-degenerate excited states. Because CIS models cannot account for static correlation stabilization of these states, transition energies to them are overestimated.

Various attempts have been made to improve CIS through perturbative corrections (CIS(D)) or incorporating select double excitations.[125,126] These approaches generally improve the representation of the ground state much more than the excited states, paradoxically leading to an even more unbalanced treatment of the ground and excited states and further divergence of computed excitation energies from the

expected values.[127]

Time-dependent density functional theories are also fundamentally based upon applying exchange-correlation corrections to CIS-like states.[128] However, because all DFT-based models are non-variational and because TD-DFT models inherit all the same limitations as the underlying CIS approach, they will not be discussed in any further detail here.

4.1.2 Excited State Coupled Cluster Theory

Analogous to CI theory, the CC equations can be extended to model excited states[129] using either a response theory[130] or equation-of-motion approach.[131] Of these, EOM-CC methods are more widely implemented and used.

As in the ground state case, EOM-CC benefits from the inclusion of higher-order excitations to give a better description of the wavefunction than an equivalent order CI expansion. Due to this, EOM-CC models provide an effective way of recovering dynamic correlation corrections and are considered the most accurate single-reference excited state model.[132] They provide high quality excited state energies and wavefunctions for single-reference systems, where HF provides an appropriate mean-field ground state model.

However, CC methods suffer from the same problems outlined above for CI methods for systems that are intrinsically multi-reference in the ground state. For strongly-correlated excited states, high levels of excitation may be required.[133] Even in the best case scenario, CC theory is based around using large numbers of determinants to generate a wavefunction that incorporates dynamic correlation, so is computationally intensive and lacks physical interpretability.

4.1.3 Multi-Configuration Self-Consistent-Field Models

As with ground state CI, the most rigorous solution to the limitations of the single reference approximation is to perform the optimization of each determinant concur-

rently - a MCSCF calculation. MCSCF methods can give very high quality excited state energies and wavefunctions - particularly if *post-hoc* corrections are applied to account for dynamic correlation - because they find optimized orbitals for the excited state determinants which provide a physically-meaningful basis for the higher-energy states. Furthermore, if an appropriate set of excitations is chosen they inherently account for any multi-reference character of the excited states, in the same manner as the ground state.[134]

However, all the problems of MCSCF calculations in the ground state remain. In fact, choosing a set of determinants for excited state calculations is often even more challenging than for the ground state because the relevant orbitals are often less well understood. A poor choice of determinants can also result in a highly unbalanced description of the excited states if a determinant relevant to the physics of some but not all of the states is excluded from the expansion.[135]

In practice these concerns are usually ameliorated through the use of active space based methods like CASSCF but, as with ground state calculations, these lead to larger CI expansions and consequently more expensive and complicated wavefunctions. Even then, selection of which and how many orbitals to include in the CASSCF active space is non-trivial, and results can vary significantly depending on the initial composition of the active space.

An additional complication unique to excited state MCSCF calculations is the lack of a single clear target for energy minimization. Minimizing energy of only the ground state will likely lead to MOs that do a poorer job of describing the excited states. The solution to this is to use some kind of state averaging, where each of the states of interest contribute some fraction of the overall energy to be minimized.[136] But exactly how this state averaging should be performed is left to the judgement of the individual user.

The requirements of selecting determinant sets and state averaging schemes mean that multi-reference methods often require the user to exercise significant expert

judgement to design calculations. The number of parameters involved in defining an MCSCF model limits systematic comparison of these methods between different systems.[137]

In principle, MCSCF approaches give us the ability to construct minimal-determinant models to describe both ground and excited state electronic configurations in a well-balanced and physically meaningful way. Unfortunately, in practice, the unoptimized nature of HF virtual orbitals makes selecting physically meaningful basis states *a priori* practically impossible.[132]

4.1.4 Reference States and Orbital Optimization

In principle, there is nothing stopping us from applying the same single determinant HF framework to single-reference excited states that works so well as a starting point for single-reference ground state calculations. Because the same physical principles apply to both ground and excited states, they should both admit mean-field solutions, which could either stand alone or be used as a basis for concise CI-type expansions. In practice, however, finding compact, physically-meaningful wavefunctions for electronically-excited states poses several complexities that are not present or are less significant in ground state calculations.

As discussed in Chapter 2, the HF procedure finds optimal mean-field MOs via an iterative energy minimisation process that most commonly converges to the global energy minimum, *i.e.* the ground state. In practice, HF calculations cannot be reliably converged to excited state solutions with any degree of specificity or selectivity. This is a serious problem, because single reference methods rely on these mean-field MOs to accurately represent the wavefunction for the target state of interest, and they are not generally available for excited states.

Even if higher-energy solutions to the HF equations could be reliably generated, excited states are often closely spaced in energy and thus are much less likely to be dominated by a single configuration than the ground state. This immediately ne-

cessitates a multi-reference method. However, for intrinsically multi-configurational excited states, no single solution to the HF equations will yield MOs that provide a good first order reference. Consequently, for many excited states no single solution to the HF equations will yield MOs that are suitable for small CI expansions; and single-determinant models will be particularly inappropriate and inaccurate.[138]

MOs calculated for a ground state reference are particularly ill-suited to cases where the electron density changes significantly between the ground and excited state; so-called charge transfer states.[139] The very large difference in character between states in these cases can mean that the virtual orbitals are wildly incorrect and even the occupied orbitals unsuitable for describing the excited configurations. This often results in large overestimations of the energies of the excited states.[122]

CIS seems like an appealing model because, despite its inaccuracy, it provides a black box way of generating simple approximations to excited states that can be be improved either through perturbation theory or through inclusion of higher excitation levels. Indeed, the analogy between CIS and HF has been remarked upon in the past,[125,140] and CIS theory has even been referred to as 'HF for excited states'.[141] The large body of work around improving CIS[142–144] demonstrates how a relatively simple well defined first order excited state wavefunction can provide a fertile starting point for further development. However, the relative lack of use of these methods suggests that the CIS wavefunction is not accurate and/or flexible enough for such experiments to be successful. This can fundamentally be traced back to the fact to the unoptimized HF virtual orbitals used in forming the CIS excited states do not provide a good representation of each state, and also that the ground state must be fundamentally single-reference in character.

Currently, the only way to overcome the limitations of single-reference excited state models is to allow concurrent optimization of molecular orbital and configuration interaction coefficients. MCSCF provides - with a skilfully selected set of configurations and/or active space specification - a very high quality model of the wavefunc-

tion, achieved by properly optimizing higher-energy MOs rather than prioritizing a single reference state. However the problems with ground state MCSCF described in Chapter 1 still apply. Reliance on user skill to effectively perform calculations is exacerbated by the complexity of the excited state manifold and the additional requirement of specifying a state averaging scheme.

4.1.5 Requirements for an Excited State Method

The end result of all these problems is the absence of any well-defined and easily calculated excited state wavefunction that provides a reasonable approximation for a wide range of cases and is amenable to systematic improvement.

In single-reference ground state calculations, this role can be filled by the HF model, although we have shown in the previous chapter that more sophisticated approaches are required for systems that are multi-reference in the ground state. In particular, the key to describing multi-reference ground state systems lies in finding appropriately optimized or semi-optimized molecular orbital sets in which to expand a minimal or near-minimal determinant configuration interaction wavefunction.

The exact same considerations apply to constructing new minimal-determinant excited state models. Therefore, it stands to reason that the same SF-NOCI approaches used to capture multi-reference effects in the ground state may be applied or adapted to describing excited states.

4.2 SF-NOCI for Excited States

The features of the SF-NOCI method that make it appropriate for modelling multi-reference effects in the ground state should apply equally to excited state calculations. In fact, given that a core goal of SF-NOCI calculations is to find more appropriate MO sets for constructing higher-energy determinants than the unoptimized virtual orbital sets produced during canonical Hartree-Fock calculations, we would expect it to provide an even greater improvement over orthogonal CI for

excited states compared to the ground state.

If the spin-flip molecular orbitals are a good approximation to the orbitals required by the corresponding "spin-unflipped" excited state determinants, then we anticipate that our SF-NOCI procedure can be used to generate physically meaningful and qualitatively correct minimal determinant excited state models.

4.2.1 Variants of SF-NOCI

Which kind of systems SF-NOCI models will be apply to, and whether they contain the minimal number of determinants required to model the system, will vary according to the configuration selection procedures outlined in Chapter 2.

CAS-SF-NOCI is essentially an approximation to CASSCF in the spin-flip orbital active space. Therefore we expect that it will be suitable in the same kind of situations as CASSCF, so long as the higher multiplicity orbitals are a good approximation to the fully optimized multi-reference MOs. This means that relatively small active spaces should be able to effectively model static correction (i.e. produce qualitatively correct wavefunctions) for a wide range of different electronic configurations and states. However, as with CASSCF, because determinants are generated as the complete set of electronic configurations possible within the defined active space, it will not necessarily represent a minimal determinant model in most cases. For example, it is likely that doubly-excited determinants are not required to describe singly-excited states at equilibrium geometries for systems that are otherwise single-reference in the ground state.

In such cases, flip-reversing SF-NOCI is more likely to provide a suitable minimal-determinant approach, particularly if the excited states themselves are single-reference in nature and the spin-flipped orbitals a good approximation to their "spin-unflipped" counterparts (orbitals optimized directly for singlet excited states). However, in cases where the single reference approximation does not apply to both the ground and excited state, the limited number of configurations in the FR-SF-NOCI wave-

function will not allow the expansion enough flexibility to form all physically-relevant states. This means we expect FR-SF-NOCI to produce qualitatively correct excitation energies at equilibrium and ground state energies at dissociation (as shown in the previous chapter) but not to reproduce accurate excited state energies at dissociation.

Perfect pairing SF-NOCI incorporates all double excitations that mix directly with the ground state. We therefore expect it will be effective in modelling "perfectly paired" doubly-excited states at equilibrium geometries. However, it will be unsuitable for modelling more common singly excited states, due to the lack of singly-excited determinants in the CI basis. Consequently, we see its potential use in excited state calculations as being limited in scope.

These later two versions of SF-NOCI both have the advantage of including fewer determinants than CASSCF and CAS-SF-NOCI like models, minimizing the number of unnecessary determinants. They also benefit from requiring less user input than CASSCF and especially general MCSCF models. PP-SF-NOCI only requires specifying the maximum number of spin flips while FR-SF-NOCI requires only that plus the highest level of excitation allowed. However, like CASSCF, there is no straightforward way – apart from "chemical intuition" to predict what spin-flip level and excitation order will be required for any given application.

4.2.2 Overall Prospects

However, unlike CASSCF or MCSCF, adding additional states to the SF-NOCI calculations is a relatively cheap process, requiring - at most - only an additional high-spin HF calculation is needed, followed by evaluation of an additional row of the CI Hamiltonian and diagonalization. This could potentially be done on an *ad hoc* basis, adding new configurations to existing SF-NOCI calculations as it becomes apparent the current determinant basis is insufficient. In contrast, the intrinsically coupled nature of the CI and MO coefficients in MCSCF-based methods means that

adding a new configuration always requires re-optimizing the MCSCF orbitals and subsequently recalculating the whole CI Hamiltonian. Further, the active space of SF-NOCI models is uniquely and completely defined by the spin-flip procedure - no user input is required to select molecular orbitals to include in the active space.

The weakness of SF-NOCI models compared to MCSCF is that the quality of the orbitals still decreases as the level of excitation increases. At higher excitation levels the high-spin states used as the HF reference will produce MO sets that differ more from those that would ideally describe the low spin state it is approximating, resulting in lower quality orbitals. This could potentially result in SF-NOCI performing worse for excited states than ground states as the first order reference for higher-energy states will not be of the same quality which could in turn cause it to overestimate excitation energies.

This also violates our requirement that all states be treated equally, as the ground state MOs are used in the same configuration they are optimized for, while all others require a spin flip from the reference configuration. Despite this, we suggest that SF-NOCI represents a useful compromise between the rigorous equality of states in MCSCF and the much more computationally cheap single reference methods. The implications of this compromise will be tested in the following chapter.

Chapter 5

Benchmarking SF-NOCI For Excited States

5.1 Introduction

One of the key features of the SF-NOCI family of methods is that they should be almost equally applicable to ground and excited states, provided that the determinant basis is appropriately tailored, and high-spin optimized reference orbitals represent a reasonable approximation to their low-spin counterparts.

However, additional complexities arise when modelling electronically excited states and, in particular, excited state potential energy curves. It is far less straightforward to determine, *a priori*, the minimal set of determinants required to qualitatively describe the wavefunction for any given electronic state at any given nuclear configuration.

Although in principle, this could simply be solved by increasing the size and completeness of the determinant basis (number of spin-flip reference states and excitations allowed within the spin-flip "active orbital" space), this is not in keeping with the original intent of this approach. Using more determinants than required to ensure an appropriate "statically correlated" reference has a number of practical

and philosophical drawbacks: it is more computationally intensive, does not cleanly separate static and dynamic correlation effects, and makes the wavefunction harder to interpret in physically meaningful terms. Furthermore, the higher the multiplicity of the reference orbital sets, the less valid the approximation that these orbitals resemble their low-spin counterparts.

Therefore, in this chapter we will continue to follow the approach of using SF-NOCI models designed for multi-reference effects in the ground states. We will asses the ability of these models to also capture multi-reference effects at equilibrium and along 1D potential energy curves for low-lying excited states.

5.2 Lithium Hydride

5.2.1 Background

The ground and low-lying singlet excited states of lithium hydride have been extensively studied, both spectroscopically,[101,145–148] and computationally[145,146,149–152] Early spectroscopic studies suggested that LiH dissociates into ion-pair fragments ($Li^+ \cdots H^-$) for all spectroscopically-observable singlet states except the ground state, in which the molecule dissociates into its constituent atoms in their ground electronic states.[101,153] However, more recent computational studies have shown that excited molecular states instead dissociate to excited-state atoms.

The first singlet excited state is of $A^1\Sigma^+$ symmetry whose equilibrium wavefunction is dominated by an electronic configuration with a doubly-occupied core Li $1s$ orbital, singly-occupied Li $2s$ + H $1s$ bonding orbital and a singly-occupied Li $2p_z$ + H $1s$ bonding orbital (assuming the molecule is aligned along the z axis). However, in this electronic state, LiH dissociates to a ground state hydrogen atom and an excited state lithium atom in its $1s^2\ 2p_z$ electronic configuration as shown in Figure 5.1

The second singlet excited state is a doubly-degenerate state of $B^1\Pi$ symmetry, with

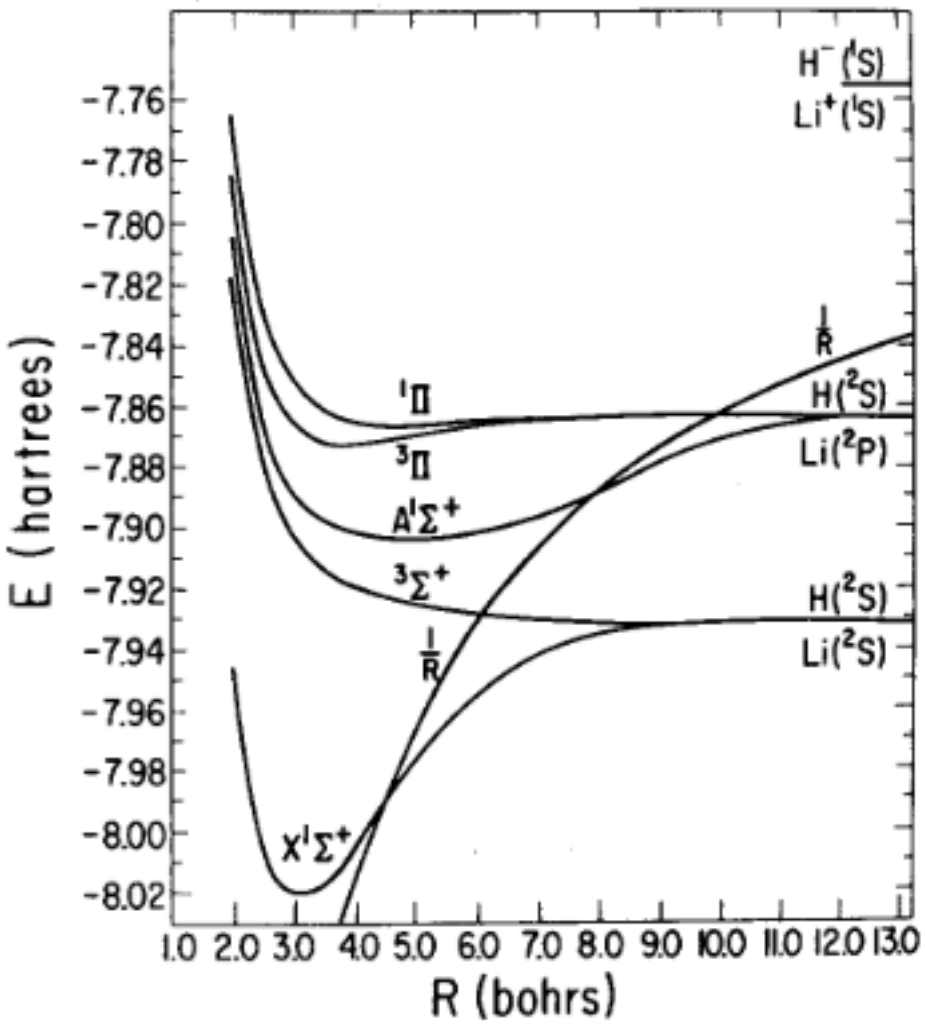

Figure 5.1: Potential energy surfaces of the first 5 states of LiH, from Docken and Hinze.[102] Also included is the Coulomb dissociation limit for $Li^+ + H^-$ - referred to as $\frac{1}{R}$

each newly occupied orbital formed as Li $2p_{x/y}$+ H $1s$ at equilibrium. However, because our SFS-NOCIS model only includes a single spin-flip reference set of orbitals, it cannot form determinants to describe this state, and so will not be considered further here.

Although the triplet states of LiH are not experimentally accessible due to spin-selection rules, they have been investigated computationally.[102,154,155] The lowest energy triplet state has been most intensively investigated and is of $a^3\Sigma^+$ symmetry. It has a purely repulsive potential energy curve and dissociates to the same energy as the $A^1\Sigma^+$ ground state, corresponding to the energy of an isolated Li atom in its ground state $1s^2\ 2s^1$ electronic configuration, plus a hydrogen atom.

5.2.2 Computational Details

The orthogonal and non-orthogonal spin-flip results described below come from the same calculations described in Chapter 3 *i.e.* SFS-NOCIS calculations in the STO-3G (for orbital analysis) and cc-pVTZ (for energies) basis sets.

CASSCF calculations were also performed - as in Chapter 3 - using an active space of 2 orbitals and 2 electrons and a single core orbital (CASSCF (2,2)). However, here the first three states are given equal weighting in forming the average density matrix in order to produce MOs that provide a better description of the higher energy states. Additionally, a second set of CASSCF calculations were carried out using a larger active space of 5 active orbitals and 2 electrons and a single core orbital, with the same averaging of states (CASSCF (5,2)).

Finally a set of CIS calculations were performed using Q-Chem, allowing single excitations into the entire virtual active space.

All the CASSCF and CIS calculations were carried out in the cc-pVTZ basis set.

5.2.3 Molecular Orbitals

All SFS-NOCIS calculations are performed in a determinant basis constructed from singlet HF orbitals (reference closed-shell determinant) and triplet CUHF orbitals (singly-excited determinants), as discussed in detail in Chapter 3. For convenience, the valence orbitals from each set at equilibrium and dissociation are reproduced in Figure 5.2.

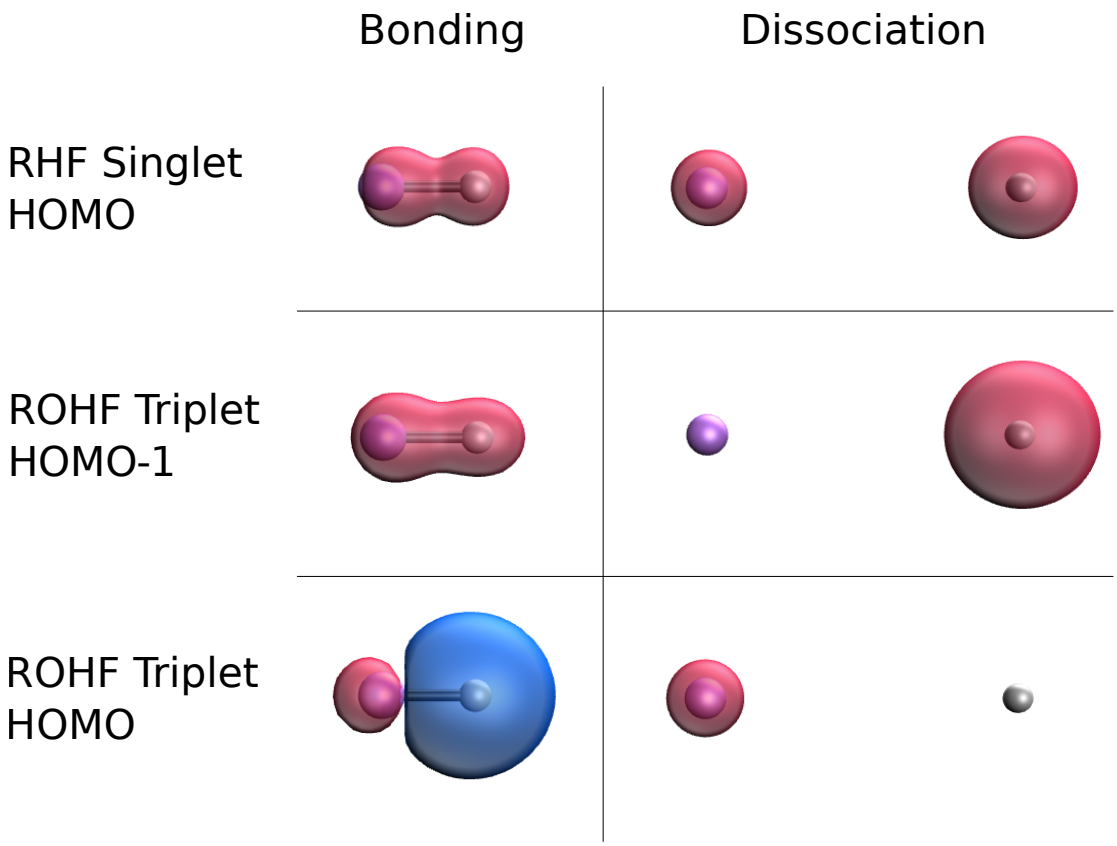

Figure 5.2: The frontier MOs of LiH at both bonding (1.6 Å) and dissociation (6.0 Å) All MOs calculated using restricted HF in the cc-pVTZ basis set. The atom on the left is H and the atom on the right is Li. In Table 5.1 RHF singlet HOMO and ROHF triplet HOMO-1 are both denoted as 2σ (despite differing in atomic composition according to spin multiplicity of the reference state) and the ROHF triplet as 3σ.

Because the wavefunction is formed from only three determinants, only two excited states can be formed in addition to the ground state. From previous experimental and computational studies, the two lowest energy excited states are expected to correspond to the lowest energy ($a^3\Sigma^+$) triplet state and the lowest energy ($A^1\Sigma^+$) singlet excited state.

The frontier orbitals of the two state averaged CASSCF calculations are visually indistinguishable from the triplet ROHF orbitals shown in Figure 5.2 indicting that the multi-reference optimization process does not qualitatively change the character of the valence orbitals, even with the additional configurations included in the CASSCF(5,2) calculation. However the HOMO+2 in the CASSCF(5,2) calculation does contribute significantly to some of the excited states. It is shown in Figure 5.3. At equilibrium this orbital is characterized by the out of phase combination of the Li p_z AO and the H $3s$ AO. As with the other molecular orbitals, it becomes almost purely atomic at dissociation, in this case forming an Li $2p_z$ orbital.

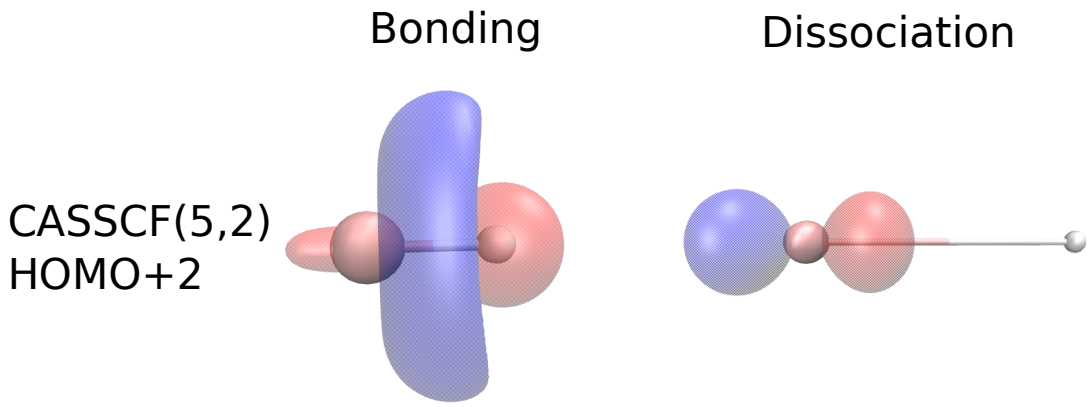

Figure 5.3: The HOMO+2 of LiH calculated using the CASSCF(5,2) level of theory, at both bonding (1.6 Å) and dissociation (6.0 Å)

5.2.4 The $a^3\Sigma^+$ State

Predicted dissociation curves for LiH in its $a^3\Sigma^+$ state are illustrated in Figure 5.4. CI coefficients for the LiH wavefunction are reported in Table 5.1, for LiH at equilibrium (1.6 Å), in the intermediate bond length regime (3.0 Å) and near the dissociation limit (4.0 Å).

For all methods other than CIS, the $a^3\Sigma^+$ state is composed of purely the out of phase combination of the two $1\sigma^2 2\sigma^1 3\sigma^1$ configurations. For most methods this is determined by symmetry as these are only two triplet configurations. The CASSCF(5,2) wavefunction contains configurations that could contribute to this state (e.g. the $1\sigma^2 2\sigma^1 4\sigma^1$ configuration that appears in the CIS wavefunction), but these configu-

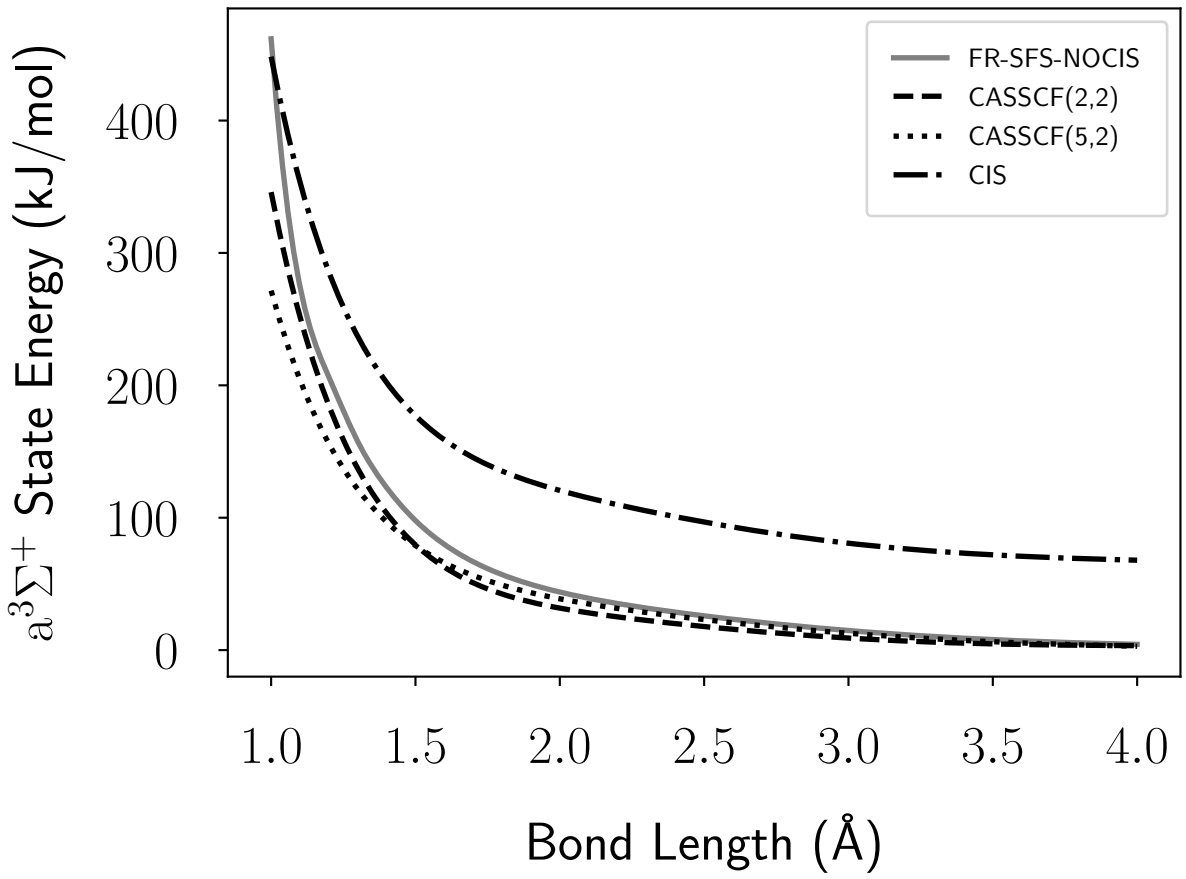

Figure 5.4: Dissociation curves of the $a^3\Sigma^+$ state of LiH at 4 different levels of theory. The energy of each curve has a been scaled so that the energy of the $X^1\Sigma^+$ state is zero. The SFS-CISD and FR-SFS-NOCIS curves have been consolidated as they are identical.

Method	Determinants	Ref	Equilibrium	Intermediate	Dissociation
FR-SFS-NOCIS	$1\sigma^2 2\sigma^1 3\sigma^1$	T	± 0.7071	± 0.7071	± 0.7071
SFS-CISD	$1\sigma^2 2\sigma^1 3\sigma^1$	T	± 0.7071	± 0.7071	± 0.7071
CASSCF(2,2)	$1\sigma^2 2\sigma^1 3\sigma^1$	$M_{(2,2)}$	± 0.7071	± 0.7071	± 0.7071
CASSCF(5,2)	$1\sigma^2 2\sigma^1 3\sigma^1$	$M_{(5,2)}$	± 0.7071	± 0.7071	± 0.7071
CIS	$1\sigma^2 2\sigma^1 3\sigma^1$	S	± 0.5983	± 0.5505	± 0.6644
	$1\sigma^2 2\sigma^1 4\sigma^1$	S	± 0.2962	± 0.3330	0.0000
	$1\sigma^2 2\sigma^1 5\sigma^1$	S	0.0000	± 0.1892	± 0.2205
	$1\sigma^2 2\sigma^1 6\sigma^1$	S	∓ 0.1744	± 0.0149	0.0000
	$1\sigma^2 2\sigma^1 7\sigma^1$	S	∓ 0.1197	± 0.1406	0.0000
	$1\sigma^2 2\sigma^1 9\sigma^1$	S	± 0.0855	0.0000	0.0000
	$1\sigma^2 2\sigma^1 12\sigma^1$	S	0.0000	0.0000	± 0.0797

Table 5.1: The non-zero CI coefficients for each for each of the methods at three different points along the dissociation curve for the $a^3\Sigma^+$ state; equilibrium (1.6 Å), intermediate (3.0 Å) and dissociation (4.0 Å). The **Ref** column specifies the reference configuration used to optimize the MOs on each row: S means a singlet reference, T means a triplet reference, $M_{(A,B)}$ means a multi-configuration optimized reference for the A orbital, B electron active space.

rations all have zero CI coefficients at all points along the dissociation curve.

Overall, this suggests that minimal-determinant models provide a reasonable representation of the $a^3\Sigma^+$ state wavefunction and are not particularly sensitive to how the underlying orbitals are obtained - whether they are optimized during a single-reference HF triplet calculation or using a state-averaged MCSCF procedure.

This is confirmed by inspection of the potential energy curve (Figure 5.4) which shows that all four models that involve some degree of orbital optimization - SFS-CISD, SFS-NOCIS, CASSCF(2,2) and CASSCF(5,2) all yield very similar dissociation curves that converge to the same dissociation limit as each other, and also the same dissociation limit as the ground state $X^1\Sigma^+$ curve.

However, the CIS curve is substantially higher in energy than the others, providing a variationally poorer solution in spite of the fact that the wavefunction comprises a much larger set of determinants than any of the other methods (Table 5.1). Clearly, including additional determinants in the CI expansion is not sufficient to compensate

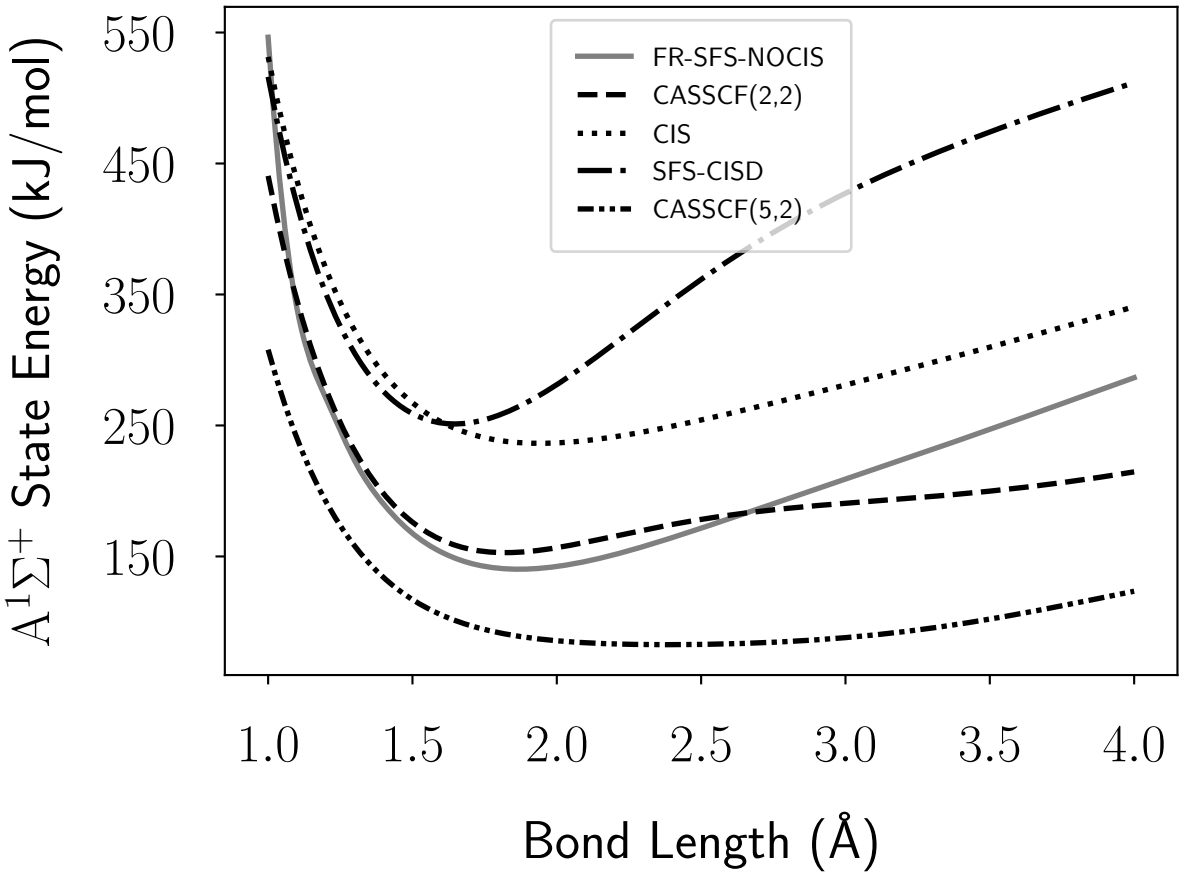

Figure 5.5: Dissociation curves of the $A^1\Sigma^+$ state of LiH at 4 different levels of theory. The energy of each curve has a been shifted so that the energy of the $X^1\Sigma^+$ state is zero.

for lack of orbital optimisation where required, *i.e.* for orbitals that electrons are excited into when forming excited states.

5.2.5 The $A^1\Sigma^+$ State

Predicted dissociation curves for LiH in its $A^1\Sigma^+$ state are illustrated in Figure 5.5 and selected CI coefficients are reported in Table 5.2.

From the predicted dissociation curves, it is immediately clear that the electronic structure of this excited state is more complicated and harder to model than the $a^3\Sigma^+$ state. It is clear that none of the curves based on small CI expansion resemble the CASSCF(5,2), indicating this state cannot be adequately modelled without the additional configurations included in that calculation. However, there are still dif-

Method	Determinants	Ref	Equilibrium	Intermediate	Dissociation
FR-SFS-NOCIS	$1\sigma^2 2\sigma^2$	S	0.4095	0.9112	1.112
	$1\sigma^2 2\sigma^1 3\sigma^1$	T	-0.7187	-0.6473	-0.533
SFS-CISD	$1\sigma^2 2\sigma^2$	T	0.6230	0.8539	0.9556
	$1\sigma^2 2\sigma^1 3\sigma^1$	T	0.5468	0.3650	0.2076
CASSCF(2,2)	$1\sigma^2 2\sigma^2$	$M_{(2,2)}$	-0.3077	0.5821	0.8605
	$1\sigma^2 2\sigma^1 3\sigma^1$	$M_{(2,2)}$	0.6708	-0.5709	-0.3592
	$1\sigma^2 3\sigma^2$	$M_{(2,2)}$	-0.0733	-0.0961	0.0392
CASSCF(5,2)	$1\sigma^2 2\sigma^2$	$M_{(5,2)}$	-0.1730	-0.4467	0.3070
	$1\sigma^2 2\sigma^1 3\sigma^1$	$M_{(5,2)}$	0.6806	0.5841	0.0000
	$1\sigma^2 3\sigma^2$	$M_{(5,2)}$	-0.0921	-0.1347	0.0000
	$1\sigma^2 2\sigma^1 4\sigma^1$	$M_{(5,2)}$	0.0684	0.1722	0.6676
	$1\sigma^2 3\sigma^1 4\sigma^1$	$M_{(5,2)}$	0.1070	0.1039	0.0000
	$1\sigma^2 4\sigma^2$	$M_{(5,2)}$	0.0000	0.1307	-0.1102
CIS	$1\sigma^2 2\sigma^1 3\sigma^1$	S	0.6727	0.6996	0.5516
	$1\sigma^2 2\sigma^1 4\sigma^1$	S	0.1443	0.0000	-0.4346
	$1\sigma^2 2\sigma^1 5\sigma^1$	S	0.1444	-0.0728	0.0000
	$1\sigma^2 2\sigma^1 6\sigma^1$	S	0.0659	0.0000	0.0000
	$1\sigma^2 2\sigma^1 7\sigma^1$	S	0.0000	0.0000	0.0527

Table 5.2: The non-zero CI coefficients for each for each of the methods at three different points along the dissociation curve for the $A^1\Sigma^+$ state; equilibrium (1.6 Å), intermediate (3.0 Å) and dissociation (4.0 Å). The **Ref** column specifies the reference configuration used to optimize the MOs on each row: S means a singlet reference, T means a triplet reference, $M_{(A,B)}$ means a multi-configuration optimized reference for the A orbital, B electron active space.

ferences between the other curves that speak to the relative qualities of the methods, which can be explained by analysing the wavefunctions in more detail.

At short bond lengths (≈ 1.6 Å), the CASSCF(5,2) wavefunction is mostly composed of configurations that are present in the smaller expansions, so the differences in energy are more likely due to small differences in orbital optimization rather than limitations in the CI expansion.

In this regime, the remaining methods fall into two groups:

- single-reference CIS and SFS-CISD methods which overestimate the energy by approximately 250 kJ/mol relative to CASSCF(5,2)
- multi-reference SFS-NOCIS and CASSCF(2,2) methods with more thoroughly optimized MOs for each contributing determinant, which overestimate the energy by approximately 150 kJ/mol

This again illustrates the limitations associated with using a single set of HF-optimized molecular orbitals to form CI determinants and wavefunctions. In this case, it appears that both singlet and triplet reference orbital sets are equally unsuitable, as both CIS (singlet reference) and SFS-CISD (triplet reference) are equally inaccurate.

At longer bond lengths, the divergence between the 5 curves is greater. As configurations with large Li p_z character come to dominate the CASSCF(5,2) wavefunction, this wavefunction is able to correctly dissociate into an excited Li atom (as shown in Figure 5.1). All the minimal-determinant methods become dominated by the $1\sigma^2 2\sigma^2$ configuration at dissociation, as they do not contain configurations with an occupied Li p_z orbital. Consequently, they are not able to dissociate to the correct atomic configuration.

CIS does much better than SF-CISD at long distances due to its inclusion of the $1\sigma^2 2\sigma^1 4\sigma^1$ configuration, it runs closer to parallel the CASSCF(5,2) curve than either of the spin flip methods do. However, CIS lacks the two double excitations

with CI coefficients > 0.1 within the CASSCF(5,2) wavefunction. In particular, the $1\sigma^2 2\sigma^2$ determinant contributes substantially to the CASSCF(5,2) wavefunction at dissociation, but is omitted from the CIS expansion entirely. This, coupled with lack of virtual orbital optimization within the reference orbital set, means that the CIS method still overestimates the energy relative to CASSCF(5,2) near dissociation compared to either SFS-NOCI or CASSCF(2,2).

The fact that SFS-NOCI outperforms the CIS wavefunction - which contains far more determinants - and runs close to the CASSCF(2,2) curve at bonding and intermediate distances, again illustrates the value of using MOs optimized for different spin states to provide an efficient basis for CI expansion.

However, the difference between CASSCF(2,2) and SFS-NOCIS and the CASSCF(5,2) curve highlights a major weakness of any minimal CI expansion approach - if the CI expansion lacks key configurations, or those configurations are poorly optimized, there is not leeway for the other configurations to adapt to fill the gap. This means that the diagonalisation of the CI matrix cannot use other - less physically meaningful - determinants to correct for missing configurations, making selecting a correct active space far more important than in methods with larger CI expansions.

Fortunately, the flexibility and low computational cost of the SFS-NOCI approach makes it relatively straightforward to fix the problem; one can either include additional excitations within the spin-flip active space, or additional high-spin reference orbital sets, or both. In the case of dissociating LiH in its $A^1\Sigma^+$ excited state, it is now clear that an additional reference spin-flip state would be required to ensure qualitatively correct dissociation, *i.e* using the SFSD-NOCIS model.

However, even the very simple SFS-NOCIS model provides qualitatively correct and physically meaningful wavefunctions for all three electronic states at equilibrium. Therefore, it provides an appropriate starting point for modelling electronic excitation energies.

Method	**Δ E ($X^1\Sigma^+ \rightarrow A^1\Sigma^+$)**
SFS-NOCIS	301.8
SF-CISD	291.1
CASSCF(2,2)	382.5
CASSCF(5,2)	317.9
CIS	392.6
Experiment[101]	349.1

Table 5.3: Excitation energies from the ground ($X^1\Sigma^+$) electronic state of LiH to $A^1\Sigma^+$ at its equilibrium bond length (1.6 Å). All energies are in kJ/mol and are calculated in the cc-pVTZ basis set.

5.2.6 $X^1\Sigma^+ \rightarrow A^1\Sigma^+$ Vertical Excitation Energies

Table 5.3 shows the vertical excitation energies from the ground state to $A^1\Sigma^+$, computed at the ground state equilibrium bond length.

Statically correlated models are expected to underestimate the experimental excitation energy, because dynamic correlation effects are expected to preferentially stabilise the ground state relative to the excited state. In the ground state, the electron distribution is more compact, so electrons can get closer to one another and so dynamic correlation plays a larger role in stabilizing the system overall.

According to this criterion, CASSCF(2,2) and CIS do not appropriately capture static correlation effects. For CASSCF(2,2) this is presumably because the doubly-occupied bonding orbital is sufficient to represent the ground state electronic structure at equilibrium, and the additional determinants present in the CASSCF(2,2) model capture dynamic correlation effects. However, for CIS, the MO basis is optimized exclusively for the ground state, which explains its preferential overstabilisation.

Only the CASSCF(5,2), SFS-CISD and SFS-NOCI models provide an appropriate statically correlated reference for subsequent dynamic correlation corrections. This will be investigated in detail in Chapter 7.

5.3 Ethylene Excited States

5.3.1 Background

As with LiH, the low lying exited states of ethylene have been widely studied.[156–159] In the literature, the first four electronic states (including the ground state) are commonly referred to as the N, T, V and Z states.[117] N is the ground state described in Chapter 3 with ${}^1A_{1g}$ symmetry.

Identically to Chapter 3, this discussion assumes the planar ethylene molecule lies in the xy plane with the C-C bond along the x axis. At 0° the molecule is completely planar and at 90° it is fully twisted (see Figure 3.7).

T is the lowest lying excited state with ${}^3B_{3u}$ symmetry and is a triplet state resulting from the π to π^* excitation.[158] The T state has not been studied extensively, but its excitation energy at 0° has been estimated from experimental data to be around 415 kJ/mol.[160]

The V state is the next lowest lying, with ${}^1B_{3u}$ symmetry, and is the corresponding singlet generated via π to π^* excitation.[158] However numerous computational studies have shown that this state mixes strongly with the π to $3d$ π Rydberg state.[157] It has an experimental excitation energy of approximately 730 kJ/mol[117] in the planar conformation.

The Z state is the highest energy of the four and is made up of a mix of the ground state and the doubly-excited $(\pi^*)^2$ configuration. Since it is a closed-shell configuration, it is also ${}^1A_{1g}$. It has not been extensively studied.

All three excited states are known to have energy minima at 90° (twisted geometry),[161,162] in contrast to the ground state, which is planar at equilibrium.

5.3.2 Computational Details

Once again the same SF-CISD calculation is used for both the ground and excited states. However, we have introduced the singly-excited states into the SFS-NOCI calculation, making it a SFS-NOCISD calculation, this was done to allow it to also capture the two low lying states with dominant single excitation character.

Additionally, a series of CASSCF(2,2) calculations were carried out with all four states given equal weighting, and a series of CIS calculations

As with the ground states a potential energy scan was carried out starting at 0° with steps in 5° increments except for the final calculation, which was performed at 88° to avoid convergence problems at 90°. All calculations were carried out in the cc-pVTZ basis set.

5.3.3 Molecular Orbitals

The MOs shown in Figure 5.6 illustrate a problem with the conventional characterisation of ethylene's MOs; generally the HOMO and HOMO+1 are simply referred to as the π and π^* orbitals. In the planar conformation, these are the MOs formed from the carbon p_z AO. However, in the twisted conformation they are composed of the p_z AO on one carbon and the p_y AO on the second carbon, such that the lobes of the MO are orientated away from the H atoms (i.e. the MO coefficient of the p_z orbital will be large on the carbon with C-H bonds in the xz and the p_y coefficient will be large on the other carbon). In the following discussion we simply refer to these orbitals as π_{2p} and π^*_{2p}; the associated orbitals with the lobes of the MOs directed towards the H atoms (*e.g.* the ROHF triplet HOMO-1) are represented with primes e.g. π'_{2p}. Higher energy MOs are denoted using higher principal quantum number labels associated with the constituent p orbitals that provide the largest contribution e.g. π_{4p}.

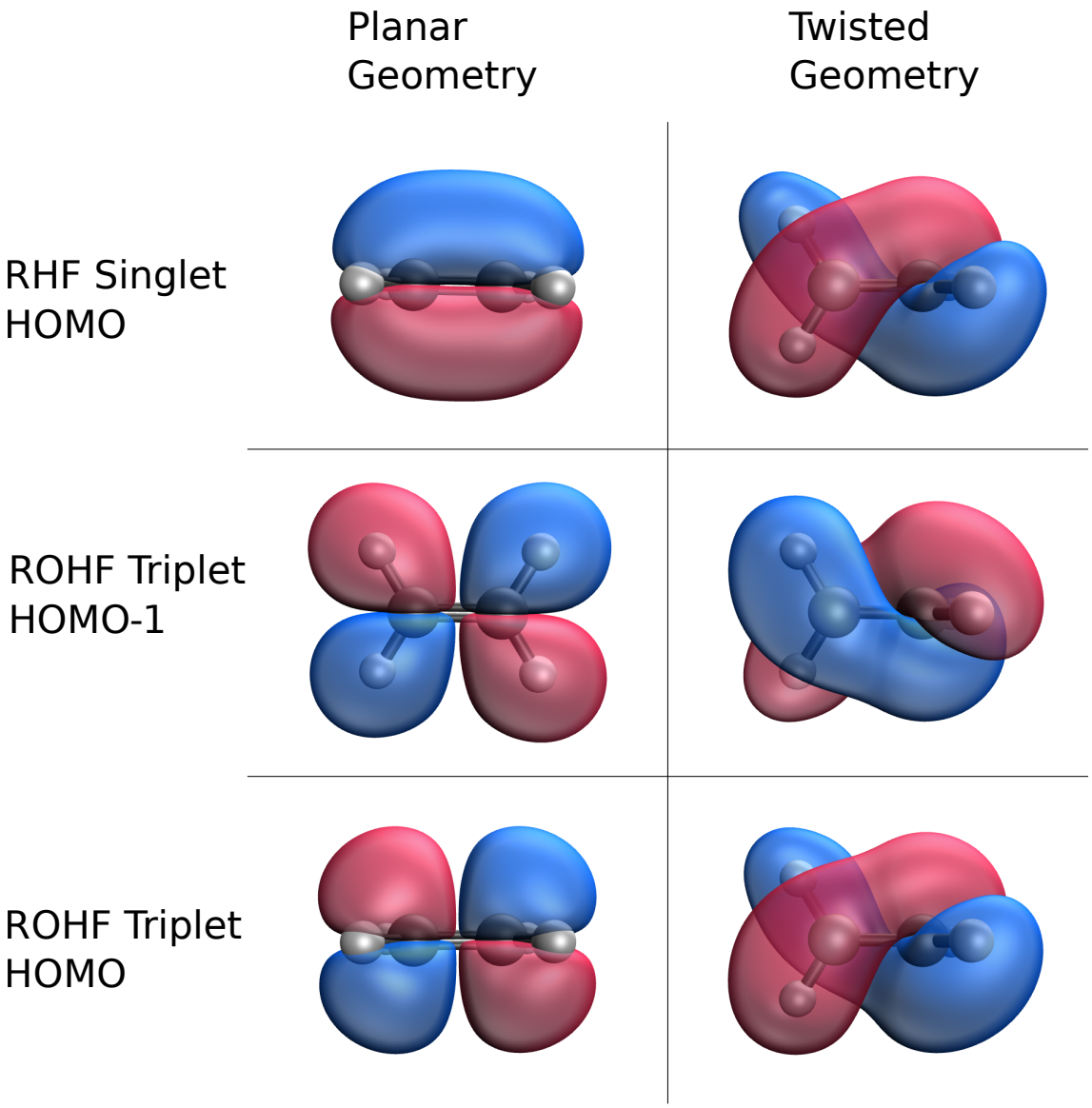

Figure 5.6: The frontier MOs of ethylene calculated in both the planar and twisted geometries in restricted HF and the cc-pVTZ basis set. Note the different orientation of the molecule in the ROHF triplet HOMO case to better display that MO. The RHF singlet HOMO and ROHF triplet HOMO-1 are the π orbital referenced in the above paragraph and the ROHF triplet HOMO is the π^* orbital.

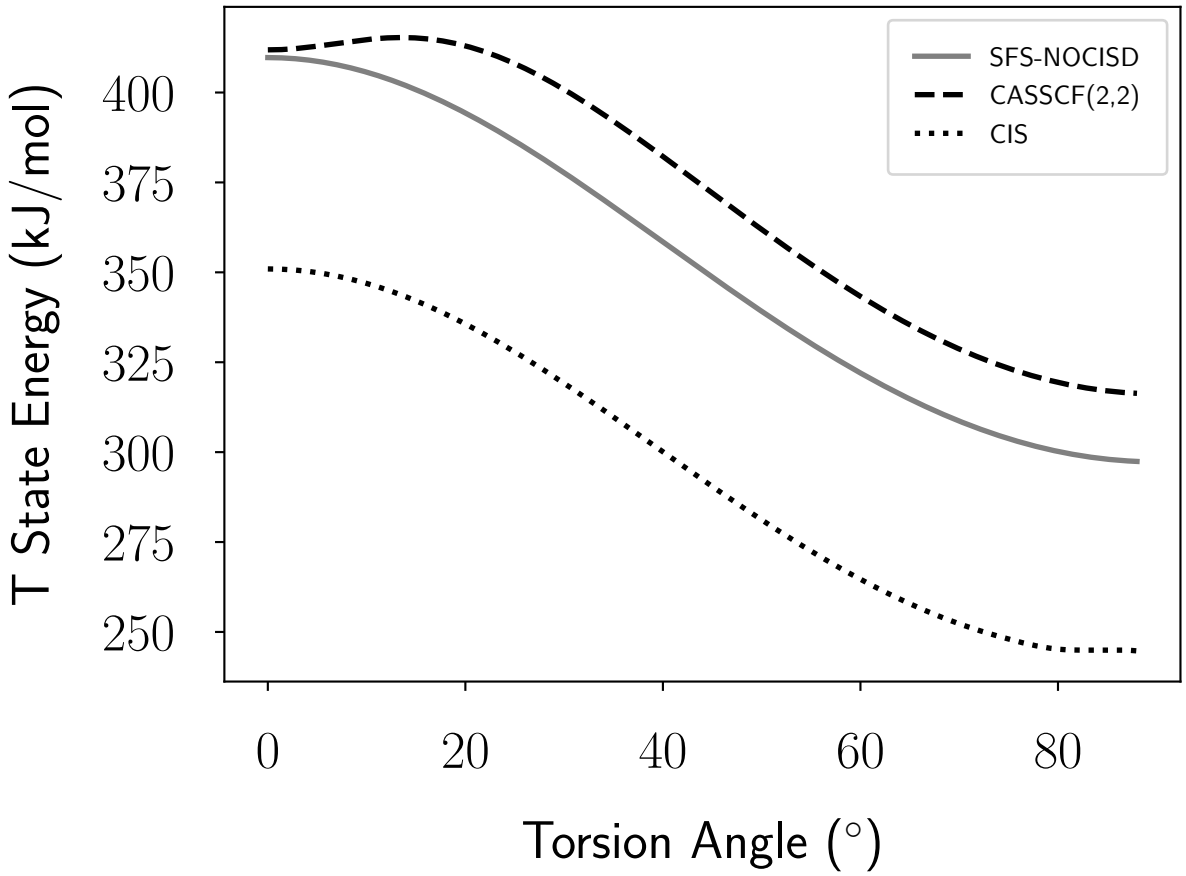

Figure 5.7: Energy of the T state (the lowest energy triplet) of ethylene as a function of H-C-C-H bond torsion angle. Each curve is shifted so that the ground state energy of the method is zero when the molecule is planar.

5.3.4 First Excited (T) State

Potential energy curves for twisting ethylene in its lowest energy triplet state are illustrated in Figure 5.7, and excitation energies relative to the ground state in the planar and twisted conformations presented in Table 5.4. CI coefficients that describe the wavefunction in these conformations are reported in Table 5.5.

The SFS-NOCISD and SF-CISD wavefunctions and energies are practically identical, because they use the same set of triplet reference orbitals to construct all determinants that contribute to the wavefunction in this state. The only difference lies in the energy shift relative to their respective ground states.

As this is a triplet state, it is forbidden by symmetry to mix with the two closed-shell configurations - meaning that the CI coefficients are completely determined by

Method	**Planar**	**Twisted**
SFS-NOCISD	410	-28
SFS-CISD	392	-9
CASSCF	411	-13
CIS	350	-233
Experiment[160]	≈ 415	–

Table 5.4: Excitation energies from the ground state of ethylene to the T state, calculated in the cc-pVTZ basis set at both 0 and 88 degrees torsion. All energies in kJ/mol

Method	**Determinants**	**Ref**	**Planar**	**Twisted**
SFS-NOCISD	$(\pi_{2p})^2$	S	0.0000	0.0000
	$(\pi_{2p})^1(\pi^*_{2p})^1$	T	$\pm$0.7071	$\pm$0.07071
SF-SDCI	$(\pi_{2p})^1(\pi^*_{2p})^1$	T	$\pm$0.7071	$\pm$0.7071
CASSCF	$(\pi_{2p})^1(\pi^*_{2p})^1$	$M_{(2,2)}$	$\pm$0.7071	$\pm$0.7071
CIS	$(\pi_{2p})^1(\pi^*_{2p})^1$	S	$\pm$0.6852	$\pm$0.6933
	$\pi_{2p} \rightarrow \pi^*_{4p}$	S	$\pm$0.1617	$\pm$0.1134

Table 5.5: CI coefficients for the T state (first excited state of ethylene). (2π) is the ground state HOMO, in the CIS case the two singly occupied orbitals are shown.The **Ref** column describes the reference configuration used to generate the MOs used in that determinant, S denotes the RHF singlet, T the CUHF triplet, $M_{(2,2)}$ the (2,2) active space MCSCF and $M_{(2,2)}$ the MCSCF containing the ground configuration plus two HOMO to LUMO single excitations. All configurations not shown are 0

symmetry with the four determinant CI basis. This directly isolates the effect of the differing optimization strategies of multi-reference vs high-spin orbitals - poor quality orbitals cannot be corrected by more mixing with other states. With minimal flexibility in the CI wavefunctions, it isn't surprising that all the potential energy curves are very close to parallel. Table 5.4 shows there is minimal difference between in the energy of the three four configuration methods - with a range of only about 5% between them.

The CASSCF energy increases to a maximum at around 15°. This maximum is only present when the MOs are optimized to provide the best average description of the first 4 states, it is not present when the T state is given significantly higher weighting in determining the MOs than the other states. This suggests that the maximum is cased by the orbitals being optimized to improve the description of the higher energy states, in a way that alters the ability of the wavefunction to describe the T state. This illustrates again how the flexibility of CASSCF models can cause incorrect results unless carefully applied. This problem is likely worsened by the MCSCF optimization needing to find orbitals that are suitable to describe both singlet (N, V, Z), and triplet (T) states - which can be avoided by using non-orthogonal orbital sets.

Aside from the artefact in the CASSCF curve, all four of the methods reproduce the rough qualitative features of the curve - a decrease in energy as the bond is twisted with the T state to be slightly lower in energy than the N state at very high torsion angles. The four configuration models also come very close to predicting the experimental planar excitation energy of 415 kJ/mol, while the CIS wavefunction significantly underestimates the energy of the T state relative to the N state at both planar and twisted configurations.

The CIS wavefunction also has only a small contribution from a single extra determinant indicating that this state is genuinely dominated by the $(\pi_{2p})^1(\pi^*_{2p})^1$ configuration and cannot be improved by mixing in other single excitations. This -

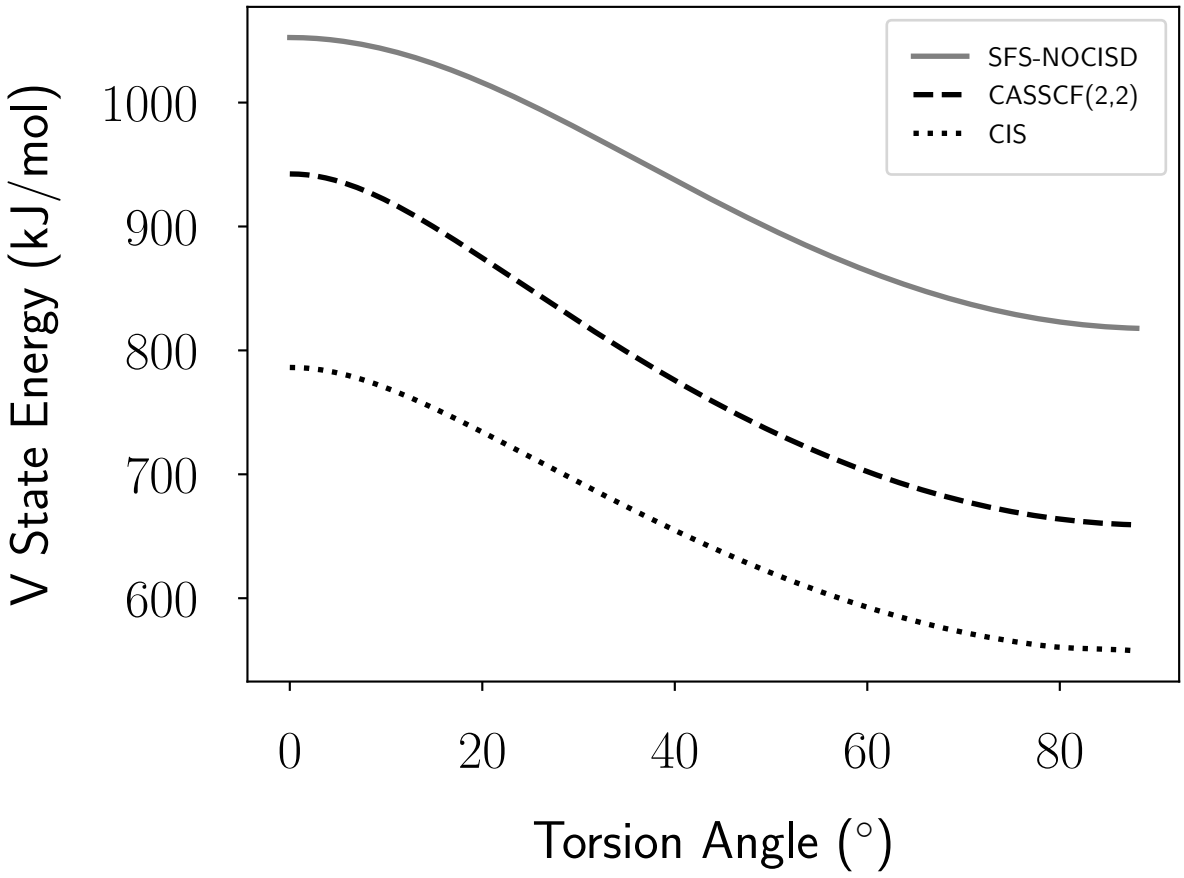

Figure 5.8: Energy of the V state ($(\pi_{2p})^1(\pi^*_{2p})^1$) of ethylene over the range of bond torsion, each curve is shifted so that the ground state energy of the method is zero when the molecule is planar.

along with the excitation energies - suggests that the four determinant model appropriately models static correlation in this case, using a small number of optimized determinants without needing to include other higher energy configurations.

5.3.5 Second Excited (V) State

Potential energy curves for twisting ethylene in its lowest energy triplet state are illustrated in Figure 5.8, and excitation energies relative to the ground state in the planar and twisted conformations presented in Table 5.6. CI coefficients that describe the wavefunction in these conformations are reported in Table 5.7.

The V state is effectively the singlet analogue of the T state and, as such, is formed from the same set of underlying triplet orbitals. Therefore, the SFS-NOCISD and

Method	**Planar**	**Twisted**
SFS-NOCISD	1052	492
SFS-CISD	1034	511
CASSCF	942	329
CIS	786	79
Experiment[117]	732	–

Table 5.6: Excitation energies from the ground state of ethylene to the V state, calculated in the cc-pVTZ basis set at both 0 and 88 degrees torsion. All energies in kJ/mol

Method	**Determinants**	**Ref**	**Planar**	**Twisted**
SFS-NOCISD	$(\pi_{2p})^2$	S	0.000	0.0006
	$(\pi_{2p})^1(\pi^*_{2p})^1$	T	0.7071	0.7071
	$(\pi^*_{2p})^2$	T	0.0000	0.0010
SFS-CISD	$(\pi_{2p})^1(\pi^*_{2p})^1$	T	0.7071	0.7071
CASSCF	$(\pi_{2p})^1(\pi^*_{2p})^1$	$M_{(2,2)}$	0.7071	0.7071
CIS	$\pi_{2p} \rightarrow \pi^*_{2p}$	S	-0.6848	0.6932
	$(\pi_{2p})' \rightarrow \pi'_{4p}$	S	-0.0710	0.0000
	$\sigma_{2p_x} \rightarrow R$	S	0.0697	0.0000
	$\sigma_{2p_x} \rightarrow R'$	S	0.0254	0.0000
	$\pi'_{2p} \rightarrow \pi^*_{2p}$	S	0.0000	0.0718
	$\pi_{2p} \rightarrow \pi^*_{4p}$	S	0.0000	0.0566

Table 5.7: CI coefficients for the V state (second excited state of ethylene). (2π) is the ground state HOMO, in the CIS case the two singly occupied orbitals are shown. R and R' are Rydberg orbitals arising from the σ^*_{2s} and $\sigma^*_{2p_x}$ MOs respectively.

SFS-CISD wavefunctions are again practically identical, and their potential energy curves overlap.

Previous computational studies have found that the V state mixes strongly with Rydberg states, and is considered notoriously difficult to model accurately.[163,164] The importance of higher energy states is hinted at by the CIS coefficients in Table 5.7, which suggest that excitations to higher energy orbitals, including Rydberg orbitals, are required. However, it is worth keeping in mind that excitations to higher energy orbitals can also be used to correct for the lack of orbital optimization within the set of HF virtual orbitals from which the CIS wavefunction is constructed.

The other three methods are much more constrained by the minimal determinant CI basis, all composed of the in phase combination of the two $(\pi_{2p})^1(\pi^*_{2p})^1$ configurations. The SFS-NOCISD wavefunction has nearly negligible contributions from the $(\pi_{2p})^2$ and $(\pi^*_{2p})^2$ configurations in the twisted conformation. However, this has only a very small effect on its energy - a change of 10.2 kJ/mol in absolute energy at 88° relative to SFS-CISD.

Of all the methods tested here, it is actually CIS that most closely replicates the experimental excitation energy at 0° of 730 kJ/mol, with all the other methods substantially overestimating - as has been found in past small configuration treatments of this state.[163] As noted above, the V state of ethylene has significant Rydberg character and cannot be adequately modelled by the $(\pi \rightarrow \pi^*)$ excitation, suggesting that the configurations containing higher energy excitations in the CIS wavefunction are necessary to correctly construct the state and are not included just to correct for poor quality reference orbitals. Lacking these higher energy configurations, the other methods are unable to stabilize the V state and so significantly overestimate its energy.

This points to a difference in MO nature that hasn't been discussed so far in this thesis. While the triplet MOs may differ in their orientation and nodal structure to the singlet orbitals, both sets of MOs are similarly contracted. This means the

Method	Planar	Twisted
SFS-NOCISD	1494	437
SFS-CISD	1487	521
CASSCF	1439	345
Experiment[165]	798	–

Table 5.8: Excitation energies from the ground state of ethylene to the Z state, calculated in the cc-pVTZ basis set at both 0 and 88 degrees torsion. All energies in kJ/mol

MOs in the minimal determinant basis are optimized to model contracted states and are not suitable for more diffuse states. Since there is so little flexibility allowed in this small basis, it is not possible to find MOs (and hence determinants) that are suitable for both kinds of states. In contrast, the much larger determinant basis of CIS is able to include configurations with Rydberg MOs allowing, the electrons to disperse much more widely.

5.3.6 Third Excited (Z) State

Of the four states considered here, the Z state is by far the least studied. Given that this state is dominated by a doubly-excited configuration, it does not occur in the CIS solution set, so only results for SFS-NOCISD, SFS-CISD and CASSCF calculations are presented here.

Potential energy curves are illustrated in Figure 5.9, and excitation energies relative to the ground state in the planar and twisted conformations presented in Table 5.8. CI coefficients that describe the wavefunction in these conformations are reported in Table 5.9.

Each of the methods produce a wavefunction which is dominated by the doubly-excited configuration at planar geometries and then gradually incorporate more of the ground-state configuration at higher torsion angles. Both the SFS-NOCISD and CASSCF wavefunction also include small contributions from the singly-excited configurations near 90°. Since these states have very large contributions from con-

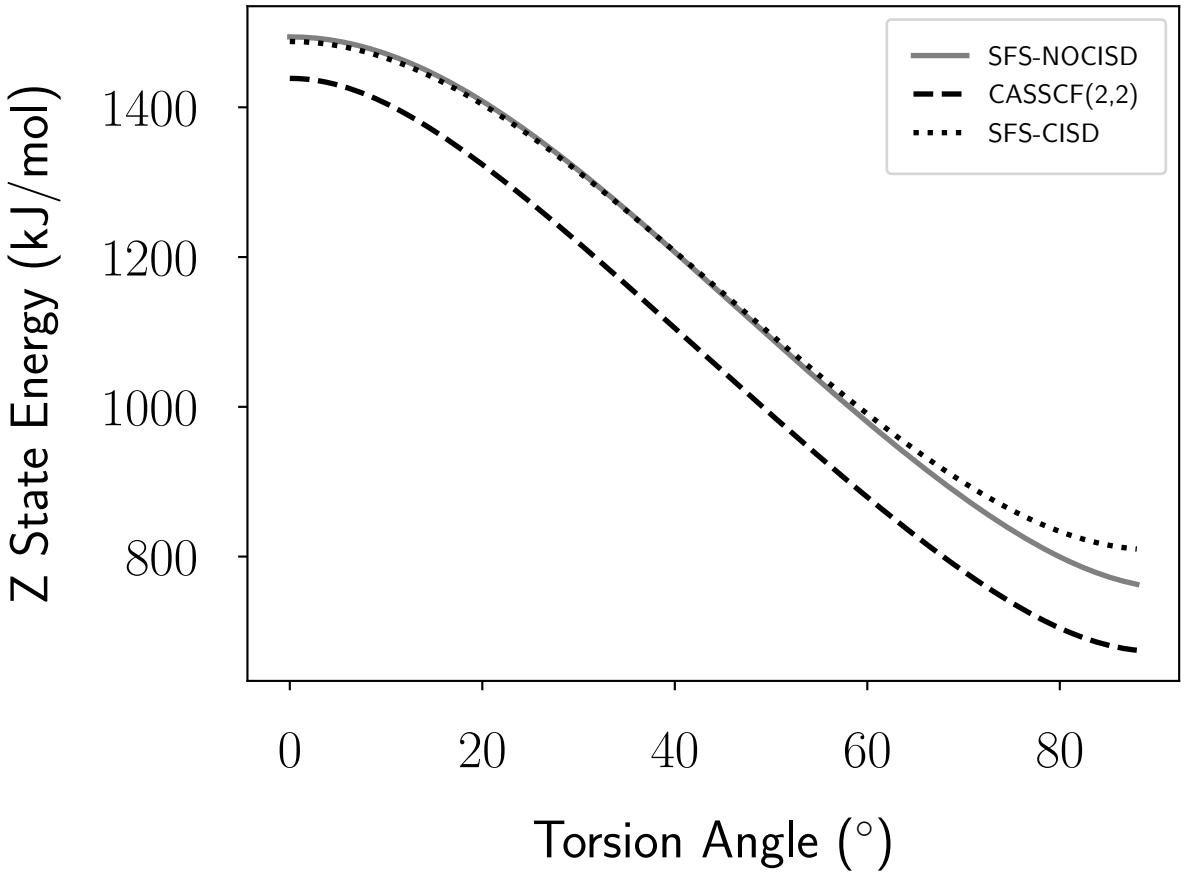

Figure 5.9: Energy of the Z state ($(\pi^*_{2p})^2$) of ethylene over the range of bond torsion, each curve is shifted so that the ground state energy of the method is zero when the molecule is planar.

Method	**Determinants**	**Ref**	**Planar**	**Twisted**
SFS-NOCISD	$(\pi_{2p})^2$	S	-0.2030	0.6340
	$(\pi_{2p})^1(\pi^*_{2p})^1$	T	0.0000	-0.0008
	$(\pi^*_{2p})^2$	T	0.9792	0.7333
SFS-CISD	$(\pi_{2p})^2$	T	0.2215	0.6792
	$(\pi^*_{2p})^2$	T	0.9752	0.7339
CASSCF	$(\pi_{2p})^2$	$M_{(2,2)}$	0.1541	0.6645
	$(\pi_{2p})^1(\pi^*_{2p})^1$	$M_{(2,2)}$	0.0000	-0.0030
	$(\pi^*_{2p})^2$	$M_{(2,2)}$	0.9880	0.7472

Table 5.9: CI coefficients for the Z state (third excited state of ethylene)

figurations with a bond order of zero, the torsional barrier is much more negative than it is for either of the primarily singly-excited states.

The energies the two spin flip methods are practically identical for most of the curve, and overestimate the excitation energy relative to CASSCF. However at approximately 60°, the SFS-CISD curve begins to level out more quickly than the SFS-NOCISD curve, overestimating the energy slightly more than the SFS-NOCISD method.

All the methods substantially overestimate the excitation energy relative to experiment (by > 600 kJ/mol). In the absence of other methods to compare to it's difficult to assess how much of this error is due to missing configurations and how much due to the exclusion of dynamic correlation corrections. This will be examined in greater detail in Chapter 7.

The excitation energies of the three methods are also all very similar in the planar configuration, but differ much more substantially in the twisted configuration. This suggests that the Z state is dominated by the single $(\pi^*_{2p})^2$ configuration at 0°. The bifurcation of the SFS-CISD and SFS-NOCISD near 90° are reflective of the increasingly multi-configurational nature of the wavefunction. At small dihedral angles the triplet based MOs of SFS-CISD can suitably model the wavefunction, but at larger angles, where both the $(\pi_{2p})^2$ and $(\pi^*_{2p})^2$ configurations are significant, either multi-reference or non-orthogonal MOs are needed.

5.4 Conclusion

Minimal determinant SF-NOCI models that are appropriate for describing ground state dissociation are also suitable for describing vertical excitation energies to low-lying excited states, provided that the system in question is primarily single-reference at equilibrium. These models also reliably capture static correlation effects in the lowest-lying triplet state potential energy curves for LiH dissociation and C-C bond rotation in ethylene. In general, where excited states can be exclusively or primarily

represented using only valence MOs, SF-NOCI produces orbitals and wavefunctions of similar quality to more expensive CASSCF models.

However, our approach of using small sets of determinants is not effective in situations where the excited states mix with configurations outside the valence space. In these cases, extended sets of determinants are required to adequately model static correlation. Fortunately, it appears that the requisite determinants could be generated by simply extending the maximum spin-flip excitation level and generating additional reference orbital sets, rather than increasing wavefunction expansion excitation level or including unoptimized virtual orbitals.

Although this approach would potentially obscure the conceptual clarity of the minimal determinant model for states that are otherwise well described by lower order methods, it would nonetheless produce a far more compact wavefunction than equivalent CASSCF and/or CIS models.

For now, however, the remainder of this thesis will focus on situations where existing models provide a good statically correlated reference model and investigate their suitability as a starting point for recovering dynamic correlation.

Chapter 6

Single and Multi-Reference Perturbation Theory

6.1 Dynamic Correlation

The SF-NOCI approaches described in the previous chapters provide an effective way to incorporate static correlation into electronic structure calculations; allowing for qualitatively-correct wavefunctions over a wide range of different geometric and bonding regimes. The use of the non orthogonal spin flipped determinants allows SF-NOCI to achieve this using a minimal number of determinants. However, because of the small number of determinants, these wavefunctions are not designed to capture dynamic correlation effects - which typically require many small contributions from many different determinants, when the CI wavefunction is expanded in a single-particle mean-field basis.[166] The previous two results chapters show that SF-NOCI models can provide qualitatively correct wavefunctions and energies, but lack quantitative accuracy due to their inability to recover the dynamic correlation contribution to the total energy.

As mentioned in Chapter 1, many body perturbation theory (MBPT) is a simple and widely used method used to recover dynamic correlation energies, post Hartree-Fock.

Perturbation theory is relatively computationally cheap and well understood. Although it cannot account for static correlation, it is an attractive post-HF correction in many cases where static correlation is negligible or has already been accounted for. In particular, second order perturbation theory (MP2), recovers a substantial proportional of the dynamic correlation energy for moderate computational cost.[23] It represents a "sweet spot" in the trade-off between accuracy and computational effort. For these reasons, second-order perturbation theory represents a promising approach to improving the quality of our SF-NOCI results.

This follows the precedent set by MCSCF-based perturbation theories, using perturbative corrections to recover dynamic correlation energies after static correlation has been accounted for by other means (such as in the widely used CASPT2 method).[59,167] It also maintains a very clear distinction between our handling of static correlation - through incorporating relevant configurations in the SF-NOCI expansion - and dynamic correlation - through the use of MBPT.

This chapter will describe the mathematics behind MP2 in both the single configuration case, and our approach to applying it to SF-NOCI wavefunctions.

6.2 Single Configuration MP2

6.2.1 Rayleigh-Schrödinger Perturbation Theory

In the standard perturbation theory approach, the exact Hamiltonian is represented as a combination of a zeroth order Hamiltonian $\hat{\mathbf{H}}^{(0)}$ (with eigenfunctions Ψ_i) plus a perturbation $\hat{\mathbf{V}}$ representing higher order corrections. The perturbation is scaled by an arbitrary parameter $0 \leq \lambda \leq 1$, which represents the degree of perturbation. This yields a modified form of the Schrödinger equation:

$$\hat{\mathbf{H}} = \hat{\mathbf{H}}^{(0)} + \lambda\hat{\mathbf{V}} \tag{6.1}$$

$$(\hat{\mathbf{H}}^{(0)} + \lambda\hat{\mathbf{V}})\Psi_i = E_i\Psi_i \tag{6.2}$$

This allows the exact energy and wavefunction to be expanded as power series.

$$E = \sum_n \lambda^n E^{(n)} \tag{6.3}$$

$$\Psi = \sum_n \lambda^n \Psi^{(n)} \tag{6.4}$$

Where $E_i^{(n)}$ and $\Psi_i^{(n)}$ are the nth order corrections to the ith eigenvalue and eigenfunction of $\hat{\mathbf{H}}$, *i.e.* to the ith quantum state and its energy. By expanding out these expressions and equating the terms for each λ^n, we can find expressions for these corrections. For full working, see Appendix B.

The zeroth order terms are just the unperturbed wavefunction and energy - the eigenfunctions and eigenvalues of the zeroth order Hamiltonian. The first order terms accounts for the effect of the perturbation operator on the 0th order wavefunction, for any given state, i. Only at second order do excited state terms start to contribute to the perturbative energy correction.

$$E_i^{(0)} = \langle \Psi_i^{(0)} | \hat{\mathbf{H}}^{(0)} | \Psi_i^{(0)} \rangle \tag{6.5}$$

$$E_i^{(1)} = \langle \Psi_i^{(0)} | \hat{\mathbf{V}} | \Psi_i^{(0)} \rangle \tag{6.6}$$

$$E_i^{(2)} = \sum_{n \neq i} \frac{| \langle \Psi_i^{(0)} | \hat{\mathbf{V}} | \Psi_n^{(0)} \rangle |^2}{E_i^{(0)} - E_n^{(0)}} \tag{6.7}$$

Although elegant, this formal result does not become useful or physically meaningful until a concrete form for both the Hamiltonian and 0th order wavefunction are specified.

6.2.2 Møller-Plesset Perturbation Theory

To apply the above formalism to the HF energy, we need to chose a concrete partitioning of the Hamiltonian into $\hat{\mathbf{H}}^{(0)}$ and the perturbation $\hat{\mathbf{V}}$.[168] In the Møller-Plesset partitioning, for $\hat{\mathbf{H}}^{(0)}$ we use the HF Hamiltonian, i.e the sum of the one electron Fock operators for each electron in the system. For the perturbation operator, we

use the difference between the exact two-electron potential and the two-electron part of the HF Hamiltonian, *i.e.* the mean-field two-electron potential.[34,169]

$$\hat{\mathbf{H}}^{(0)} = \sum_i \hat{\mathbf{F}}_i \tag{6.8}$$

$$\hat{\mathbf{V}} = \sum_{i<j} r_{ij}^{-1} - \sum_i V_i^{HF} \tag{6.9}$$

Using this partitioning, $E^{(0)}$ and $E^{(1)}$ are the one- and two-electron parts of the HF energy respectively. Therefore $E^{(2)}$ is the first term that contributes to the correlation energy. As in CI, we can use virtual HF orbitals to construct approximations to the higher energy configurations required to evaluate equation 6.7. Since $\hat{\mathbf{V}}$ is a two-electron operator, only the terms of equation 6.7 involving double excitations from the reference determinant will be non-zero. The resulting expression for the second order MP correction is (for full details of working see Appendix B):

$$E^{(2)} = \sum_{i,j,a,b} \langle \Phi | \hat{V} | \Phi_{i,j}^{a,b} \rangle \frac{\langle \Phi_{i,j}^{a,b} | \hat{V} | \Phi \rangle}{E_0 - E_{i,j}^{a,b}} \tag{6.10}$$

$$= \sum_{i<j,a<b} \frac{|\langle \Phi | \hat{V} | \Phi_{i,j}^{a,b} \rangle|^2}{E_0 - E_{i,j}^{a,b}} \tag{6.11}$$

This expression can be reduced further by evaluating the two electron integrals using the Slater-Condon rules, and expanding the state energies in terms of molecular orbital energies:

$$E^{(2)} = \sum_{i,j,a,b} \frac{|\langle ia||jb\rangle|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \tag{6.12}$$

Here i and j index the occupied MOs, and a and b run over the virtual orbitals. ϵ_n is the energy of the nth MO. The matrix element $\langle ia||jb\rangle$ is the antisymmetrized 2 electron integral (as in Chapter 2) over the MO spin orbitals χ_i, χ_a, χ_j, and χ_b.

$$\begin{aligned}\langle ia||jb\rangle = &\int \chi_i^*(1)\chi_a^*(2)\, r_{12}^{-1}\, \chi_j^*(1)\chi_b^*(2)\mathrm{d}\mathbf{x}_1\mathrm{d}\mathbf{x}_2 - \\ &\int \chi_i^*(1)\chi_a^*(2)\, r_{12}^{-1}\, \chi_b^*(1)\chi_j^*(2)\mathrm{d}\mathbf{x}_1\mathrm{d}\mathbf{x}_2\end{aligned} \tag{6.13}$$

Evaluating this expression involves contracting atomic Coulomb integrals to form these molecular integrals. Full details of how this transformation is achieved are reported in Appendix B. Overall, the process scales as $O(N_{\text{orbs}}^5)$.

6.3 Multi-Configuration MBPT2

The above discussion only treats perturbation theory applied to correct the energy of a single HF determinant. To apply MBPT to SF-NOCI it is obviously necessary to extend this to apply to CI type wavefunctions. A number of different approaches have been developed, differing in how perturbation theory is applied to various components of the CI wavefunction, and in what order. Broadly, these approaches can be grouped into "perturb-then-diagonalize" and "diagonalize-then-perturb" methods.[170] The most commonly used multi-configurational perturbation model is CASPT2, which is a diagonalize-then-perturb method based on a CASSCF wavefunction. However, in applications of MBPT to regular CI-type wavefunctions, perturb-then-diagonalize methods seem more commonly used - perhaps to avoid problems with size-extensivity.[171]

However, additional complications arise when attempting to apply perturbation theory corrections to NOCI-based models, because many of the simplifications afforded by working in an orthogonal MO basis and orthogonal set of CI determinants no longer apply. The starting point for developing a non-orthogonal version of MRMP2 is to expand all determinants as separate Taylor series in the perturbation parameter, λ.

Following the same process used to derive the Möller-Plesset energy expression yields the following second-order correction term for two non-orthogonal determinants, Φ and $\widetilde{\Phi}$:

$$E^{(2)} = \sum_{i,a} \frac{\langle \Phi | \hat{V} | \widetilde{\Phi}_i^a \rangle \langle \Phi_i^a | \hat{V} | \widetilde{\Phi} \rangle}{(E_0 - E_i^a)} + \frac{1}{4} \sum_{i,j,a,b} \frac{\langle \Phi | \hat{V} | \widetilde{\Phi} i, j^{a,b} \rangle \langle \Phi_{i,j}^{a,b} | \hat{V} | \widetilde{\Phi} \rangle}{(E_0 - E_{i,j}^{a,b})} \tag{6.14}$$

where the subscripts i, j index occupied orbitals that electrons are drawn from and a, b the virtual MOs to which they are promoted. For full details of the derivation, see Appendix B.

The fact that the first term is not exactly zero arises from the non-orthogonality of the underlying determinants. However, in practice this term is likely to be very small and can probably be neglected.

The second term is clearly analogous to the single-determinant MP2 energy expression, equation (6.11), except that it involves evaluating electron repulsion integrals between one reference determinant and doubly-excited states of the other reference determinant. Computationally, this poses a challenge, because the occupied and virtual orbital sets cannot be mutually biorthogonalized.[172] Therefore, a separate bi-orthogonalisation process must be carried out for each doubly-excited determinant. Worse, should the bi-orthogonalised MO sets contain two pairs of MOs with non-zero overlaps, the MO integrals must then be evaluated using the $O(N_{\text{orbs}}^5)$ AO to MO transformation procedure. However, computational savings may be possible by pre-computing half-transformed integrals for given sets of orbitals and only performing the final contraction step as required.

Another alternative may be to reformulate these expressions into a more computationally manageable form, introducing resolutions of the identity to abstract the non-orthogonality problem away from the two-electron repulsion integrals into sets of overlap integrals. This approach has been pursued by Yost *et al.*[173] in the context of developing ΔSCF-based perturbation theory models, but the resultant expressions are still quite complex.

6.4 Considerations for SF-NOCI-MP2

In light of the above considerations, we propose a somewhat crude form of NOCI-MP2, in which the MP2 corrections are applied to only the diagonal elements of the CI matrix. This means that the MP2 part of NOCI-MP2 is equivalent to ap-

plying MP2 to each of the HF determinants forming the CI basis. By excluding the off diagonal terms we avoid the need to calculate expensive biorthogonal MP2 corrections as well as reduce the sheer number of corrections required from n^2 to n. This simplification has been suggested before in CI expansions based on VB-like orbitals,[174] and also in the context of enabling MRMP2 models to be applied to large molecules.[175]

To justify this simplification, we note that the off diagonal elements of the CI Hamiltonian represent the mixing between the determinants required to correct for the inability of any single one of them to adequately describe the wavefunction. Therefore the better quality the determinants the less important the off-diagonal terms will be to determining the final CI energies. The point of using optimized orbitals in SF-NOCI is precisely to provide a good quality single determinant first order wavefunction with minimal contributions from other states. This means that the influence of state mixing on the wavefunction should be smaller than in single-reference CI models, and consequently that both the off diagonal CI elements and their MP2 corrections will be smaller. Since the off diagonal elements are likely to have small contributions to the wavefunction, we believe that it is appropriate to focus on the more easily calculated diagonal terms that dominate the SF-NOCI wavefunction.

Further, it is easy to detect *a priori* where this assumption breaks down, by inspection of the CI matrix elements and the CI mixing coefficients. If the off-diagonal elements are small, and the CI coefficient vectors dominated by a single leading term, then this approach should work well.

This is framed as a perturb-then-diagonalize approach however a similar diagonalize-then-perturb method could be applied by calculating the MP2 correction to each HF state and then applying those corrections directly to the final SF-NOCI energies weighted by the contribution of each HF state.

Chapter 7

SF-NOCI-PT2 Results for Ground and Excited States

7.1 Introduction

In Chapters 3 and 5, we demonstrated that the SF-NOCI approach provides qualitatively correct wavefunctions for capturing static correlation effects in ground state dissociation curves and vertical excitation energies at equilibrium. However, as a minimal determinant model, it is intrinsically ill-suited to capturing dynamic correlation effects by design.

However, it is well known that capturing dynamic correlation is important to achieve chemical accuracy.[18] In this chapter, we will assess the ability of perturbatively corrected SF-NOCI models to recover total (statically + dynamically correlated) energies, and benchmark their performance against "conventional" multi-reference MP2 approaches, and also highly accurate single-reference models – CCSD(T) and EOM-CCSD(T) – where applicable.

7.2 Computational Details

SF-NOCI-PT2 calculations were carried out using the approach described in the previous chapter in which only the diagonal elements of the SF-NOCI Hamiltonian are corrected in a perturb-then-diagonalize framework.

MRPT2 calculations were carried out using multi-configurational quasidegenerate perturbation theory (MCQDPT) as implemented in GAMESS.[176] This is a perturb-then-diagonalize approach. A full explanation of the relationship between perturbatively-corrected multi-reference methods is provided in Appendix B.

The underlying CASSCF orbitals may be optimized either for the ground state, or averaged over the ground state and excited states of interest. The former will be denoted CASSCF and the latter CASSCF*, or MRPT2 and MRPT2* for the corresponding perturbatively corrected models.

For the dissociation of LiH and C-C bond rotation in ethylene, potential energy scans were carried out using the same approach described in Chapters 3 and 5. In brief, the LiH potential energy curve is mapped out in 0.1 Å increments from 1 Å to 4 Å, while the torsional potential energy curve is generated using a step size of 5 °, except near perpendicular where the torsion angle is set to 88° rather than 90° to avoid numerical instability.

SF-NOCI-PT2 and MRPT2 vertical excitation energies were also obtained during this process, from calculations performed at each equilibrium geometry: R(Li-H) = 1.6 Å, ϕ(HCCH) = 0°. Additionally, single point EOM-CC calculations were carried out using the CFOUR package[177] to obtain high-quality, independent excitation energy estimates for both LiH (EOM-CCSD(T)) and ethylene (EOM-CCSD). Experimental and computational reference values are also compiled from literature sources, where available.

All calculations were performed in the cc-pVTZ basis set.

All energies presented in this chapter are decomposed into contributions from stati-

cally and dynamically correlated components. Dynamic correlation components are obtained as the difference between results obtained using fully correlated models – MCQDPT, SF-NOCI-PT2 and CCSD(T) – and their statically-correlated counterparts – CASSCF, SF-NOCI and RHF.

7.3 Lithium Hydride

As demonstrated in Chapters 3 and 5, the SFS-NOCIS model provides an appropriately statically-correlated reference for the ground-state dissociation curve of LiH and vertical excitation energies to the lowest-lying singlet and triplet excited states, labelled $A^1\Sigma^+$ and $a^3\Sigma^+$, respectively.

However, while these models provide qualitatively correct wavefunctions, the predicted energies are not quantitatively accurate due to lack of dynamic correlation. This section will evaluate the ability of perturbatively-corrected models to recover the missing dynamic correlation energy.

7.3.1 Ground State ($X^1\Sigma^+$) Potential Energy Curve

Potential energy curves for dissociating LiH are illustrated in Figure 7.1, according to our statically correlated reference model (SFS-NOCIS), a series of fully-correlated models (SFS-NOCIS-PT2, MCQDPT2 and CCSD(T)) and experimental reference values extracted from vibrational spectroscopy measurements.[101] In all cases, energies are computed and reported relative to sum of the energies of the individual atoms according to each model.

From Figure 7.1 it is immediately clear that CCSD(T) most closely reproduces the experimental dissociation curve. This makes sense, because the CCSD(T) wavefunction is formed from a much more extensive determinant basis and the total energy is variationally optimized.

However, both the MCQDPT2 and FR-SFS-NOCIS-PT2 models recover the bulk

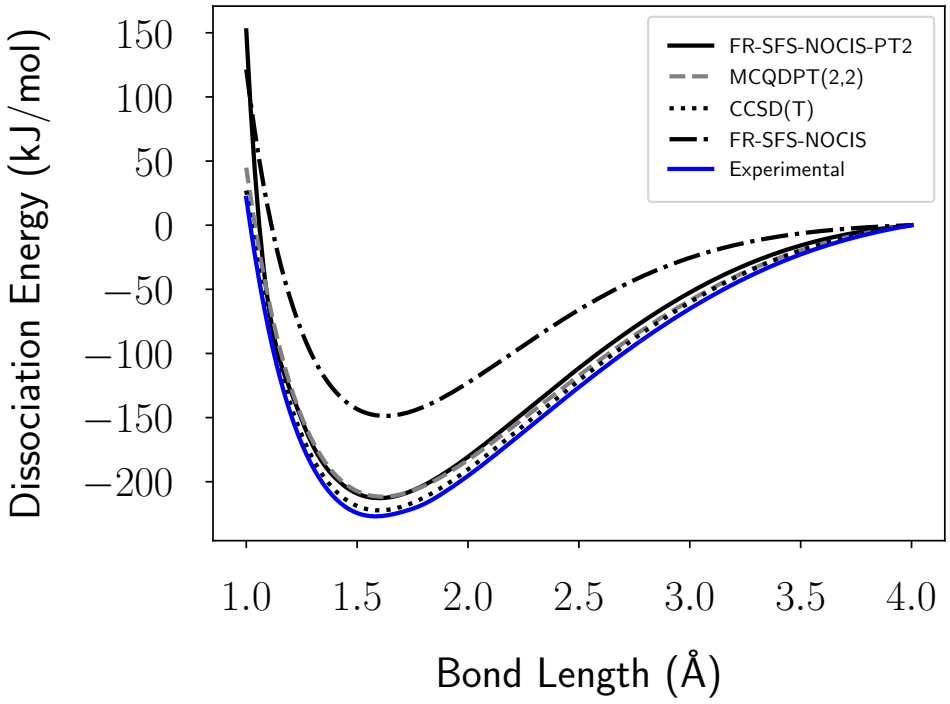

Figure 7.1: The potential energy surface of dissociating LiH calculated using, SFS-NOCIS - both with and without our perturbation correction - compared with the MCQDPT, CCSD(T) and experimental surfaces.[101] All energies are scaled to be zero at dissociation

of the dynamic correlation component of the binding energy. The potential energy curves overlap near equilibrium, but diverge slightly at stretched bond lengths, where the MCQDPT2 curve is lower in energy than the FR-SFS-NOCIS-PT2 curve, indicating a variationally superior model. This can be attributed to the fact that neither the singlet or triplet reference orbitals used in constructing the SFS-NOCIS wavefunction are ideally suited here; the singlet reference state wavefunction contains artefactual ionic terms that bias the shape of the orbitals, while the triplet reference orbitals include the effects of Fermi repulsion between opposite-spin electrons, although these obviously decay with increasing interatomic and interelectronic distance.

Nonetheless, all fully correlated models yield potential energy curves in much closer agreement to experiment than the statically-correlated FR-SFS-NOCIS model. This can be neatly summarised by computing and comparing dissociation energies predicted by each model, and referenced to experiment.

Method	D_e^{stat}	D_e^{dyn}	D_e
FR-SFS-NOCIS/+PT2	148.5	64.3	212.8
CASSCF/MCQDPT2	176.1	35.5	211.6
RHF/CCSD(T)	139.9	90.7	230.6
Experiment[101]			227

Table 7.1: Ground state dissociation energies of lithium hydride and their static and dynamic correlation contributions, calculated using matched statically and fully correlated models. A (2,2) active space is used in CASSCF and MCQDPT2 calculations. All calculations are performed in the cc-pVTZ basis set. All energies in kJ/mol.

7.3.2 Ground State ($X^1\Sigma^+$) Dissociation Energy

Dissociation energies are computed as:

$$D_e = E(R = 4.0) - E(R = 1.6) \tag{7.1}$$

When obtained using statically correlated models, a superscript notation is used: D_e^{stat}.

The dynamic correlation contribution to the overall dissociation energy is computed as:

$$D_e^{\text{dyn}} = D_e - D_e^{\text{stat}} \tag{7.2}$$

As observed qualitatively in the dissociation curves, inclusion of dynamic correlation increases the dissociation energy by preferentially stabilizing the bound state. This is as expected because the bound molecule has electrons in closer proximity than the separated atoms, hence it will have greater dynamic correlation energy.

The dynamic correlation correction to the FR-SFS-NOCIS dissociation energy is greater than its CASSCF counterpart. This is most likely due to the additional doubly-excited determinant in the CASSCF(2,2) expansion contributing to recover-

Method	$E_{\text{trans}}^{\text{stat}}$	$E_{\text{trans}}^{\text{dyn}}$	E_{trans}
FR-SFS-NOCIS/+PT2	228.4	59.6	287.9
CASSCF*/MCQDPT2*	282.5	35.3	317.8
CCSD(T)	–	–	309.8
FC LSE[155]	–	–	317.1

Table 7.2: Excitation energies for the $X^1\Sigma^+ \rightarrow a^3\Sigma^+$ transition of lithium hydride, calculated using FR-SFS-NOCIS and CASSCF wavefunctions both with and without second order perturbation theory corrections applied. A (2,2) active space is used in CASSCF and MCQDPT2 calculations. Because there is no un-correlated analogue of triplet-state CCSD(T) calculations, the total transition energy cannot be decomposed into static and dynamic correlation contributions. All calculations are performed in the cc-pVTZ basis set, except the reference free-complement local Schrödinger equation (FC LSE) results that effectively correspond to full CI calculations in a near-complete basis.[155] All energies in kJ/mol.

ing some dynamic correlation at equilibrium, but could also be due to the greater flexibility of the CASSCF orbitals. Regardless, it is clear that there is more dynamic correlation energy to be captured from a FR-SFS-NOCIS reference starting point, compared to CASSCF(2,2).

7.3.3 $X^1\Sigma^+ \rightarrow a^3\Sigma^+$ Excitation Energy

Transition energies are computed as:

$$E_{\text{trans}} = E(a^3\Sigma^+) - E(X^1\Sigma^+) \tag{7.3}$$

Values computed using statically correlated models are denoted $E_{\text{trans}}^{\text{stat}}$ and the dynamic correlation contribution computed as the difference:

$$E_{\text{trans}}^{\text{dyn}} = E_{\text{trans}} - E_{\text{trans}}^{\text{stat}} \tag{7.4}$$

As Table 7.2 shows, perturbatively accounting for dynamic correlation increases excitation energies for both the FR-SFS-NOCIS and CASSCF*(2,2) wavefunctions. Again, this is an unsurprising outcome because ground states typically have a denser electron distribution than the excited states and consequently are less well approx-

Method	$E_{\text{trans}}^{\text{stat}}$	$E_{\text{trans}}^{\text{dyn}}$	E_{trans}
FR-SFS-NOCIS/+PT2	301.8	44.6	346.5
CASSCF*/MCQDPT2*	382.5	-33.2	349.3
EOM-CCSD(T)	–	–	. 342.8
FC LSE[155]	–	–	352.1
Experiment[101]			349

Table 7.3: Excitation energies for the $X^1\Sigma^+ \rightarrow a^3\Sigma^+$ transition of lithium hydride, calculated using FR-SFS-NOCIS and CASSCF wavefunctions both with and without second order perturbation theory corrections applied. A (2,2) active space is used in CASSCF and MCQDPT2 calculations. Because there is no un-correlated analogue of EOM-CCSD(T) the total transition energy cannot be decomposed into static and dynamic correlation contributions. All calculations are performed in the cc-pVTZ basis set, except the reference free-complement local Schrödinger equation (FC LSE) results that effectively correspond to full CI calculations in a near-complete basis.[155] All energies in kJ/mol.

imated using mean-field models and so have stronger dynamic correlation energies. Therefore, the perturbation theory correction preferentially stabilizes the ground state and therefore increases the transition energy.

All methods that account for dynamic correlation come very close to replicating the benchmark theoretical excitation energy.[155] In particular, the MCDQPT2* results are very close to the reference values, although this may be due in part to error cancellation between basis set incompleteness and wavefunction truncation errors. Experimental transition energies cannot be measured for this transition, because they are spin-forbidden.

7.3.4 $X^1\Sigma^+ \rightarrow A^1\Sigma^+$ Excitation Energy

Predicted and literature reference excitation energies for the $X^1\Sigma^+ \rightarrow A^1\Sigma^+$ transition are presented in Table 7.3, decomposed into static and dynamic correlation components as described in the previous section.

Here, the dynamic correlation correction increases the transition energy for FR-SFS-NOCIS by preferentially stabilizing the ground state, as expected.

However, for CASSCF*, the perturbation theory correction decreases the transition

energy, indicating that it is stabilizing the excited state more than the ground state. This arises from the fact that the CASSCF(2,2) wavefunction already captures some dynamic correlation through inclusion of a doubly-excited determinant that mixes with the doubly-occupied bonding state, while the excited state is represented using only singly-excited determinants that can mix to ensure correct spin symmetry but not capture dynamic correlation effects. Therefore, there is effectively less dynamic correlation energy remaining to recover in the ground state, compared to the excited state.

Once again all methods that account for dynamic correlation come very close to replicating the theoretical and experimental reference excitation energies.[101,155]

7.4 Ethylene

As demonstrated in Chapter 3, the SFS-NOCID model yields qualitatively correct torsional potential energy curves for rotation about the C-C bond in ethylene, avoiding the artificial cusp at 90°. However, torsional barrier heights cannot be precisely estimated using this model, due to its inability to capture dynamic correlation effects. Here we benchmark the performance of its dynamically-correlated analogue SFS-NOCISD-PT2 against MCQDPT2.

Chapter 5 shows that SFS-NOCISD contains the dominant terms required to describe the T and Z excited state wavefunctions of ethylene. However, we did not attempt to assess the accuracy of the computed N $\rightarrow$ T and N $\rightarrow$ Z transition energies, because this model lacks the flexibility required to capture dynamic correlation effects and so is not expected to yield accurate results. Benchmarking against experimental or higher-level *ab initio* reference values both the statically-correlated SFS-NOCISD model and its dynamically correlated counterpart SFS-NOCISD-PT2 will be performed here. As with LiH, we will not attempt to apply perturbation theory corrections to excitation energies where the statically correlated wavefunctions and energies for the target state are obviously qualitatively incorrect - as is the case

for the N $\rightarrow$ V excitation.

7.4.1 Torsional Potential & Barrier Height

Calculated torsional potential energy curves are illustrated in Figure 7.2 and corresponding barrier heights tabulated in Table 7.4. Torsional potential energies are computed as:

$$E_{\text{tors}} = E(\phi) - E(0) \tag{7.5}$$

where E(0) is the energy of the molecule at it's equilibrium geometry. When computed using statically correlated models, this quantity is denoted $E_{\text{tors}}^{\text{stat}}$. Dynamic contributions to torsion energies are defined as:

$$E_{\text{tors}}^{\text{dyn}} = E_{\text{tors}} - E_{\text{tors}}^{\text{stat}} \tag{7.6}$$

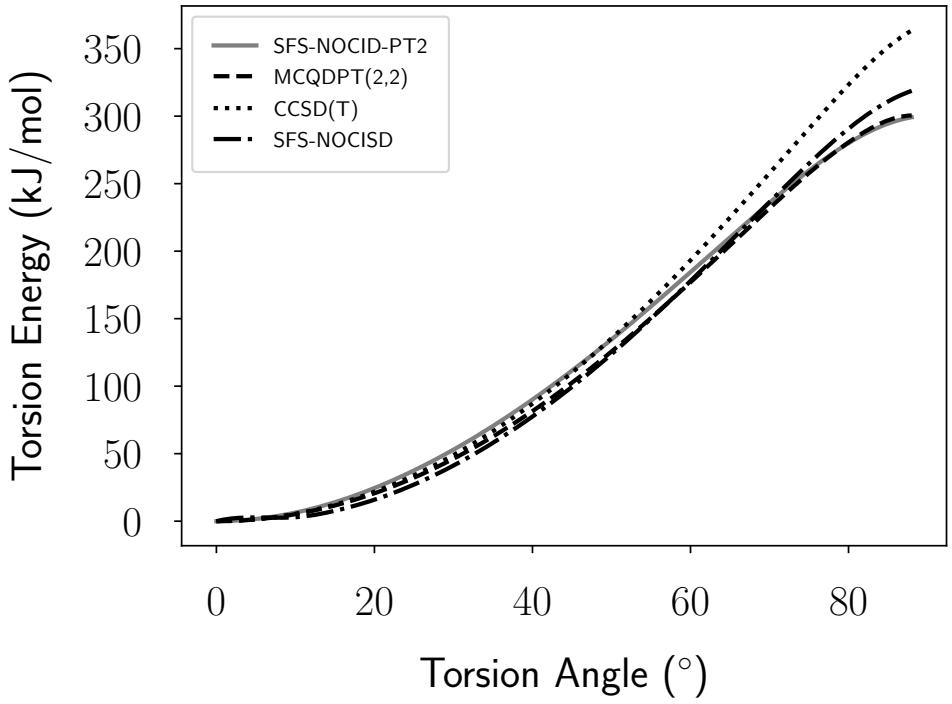

Figure 7.2: The potential energy surface of twisting ethylene calculated using, SFS-NOCID - both with and without our perturbation correction - compared with the MCQDPT and CCSD(T) surfaces. All energies are shifted to be zero at equilibrium.

From Figure, 7.2 it is evident that the single-reference CCSD(T) curve is quali-

Method	$E_{\text{tors}}^{\text{stat}}$	$E_{\text{tors}}^{\text{dyn}}$	E_{tors}
SFS-NOCID/+PT2	325.5	-26.3	299.2
CASSCF/MCQDPT2	316.4	-15.9	300.5
RHF/CCSD(T)	439.8	-74.5	365.3
MR-CISD+Q			289.5
DMC			293.7
VMC			300.8
Experiment[178]			≈ 270

Table 7.4: Torsion barriers for the ground state of ethylene calculated using the SFS-NOCID, CASSCF(2,2) and RHF wavefunctions and dynamically-correlated analogues. All calculations are in the cc-pVTZ basis set. The MR-CISD+Q, DMC and VMC calculations are from Zen *et al.*[179] All energies in kJ/mol.

tatively different to the others. However, it is known that CCSD(T) calculations cannot adequately capture static correlation effects near the barrier.[24,66] Even the simple statically-correlated SFS-NOCISD model agrees better with the SFS-NOCID-PT2 and MCQDPT2 models that are designed to capture both static and dynamic correlation effects. Clearly, appropriately modelling static correlation is required for even qualitative accuracy in this situation. However, establishing the quantitative accuracy of the multi-reference models is harder. The simplest and clearest way is to benchmark torsional barriers against experimental data or higher-level *ab initio* results.

The torsion barrier height of ethylene is defined here as the difference in energy between a torsion angle of 0° and 88°, with all bond lengths and other bond angles kept constant. Only allowing the torsion angle to reach 88° will obviously slightly underestimate the "true" total torsion barrier, however this difference should not affect the comparison between the different models. Furthermore - as shown in Figure 7.2 the PES curves for the multi-reference models all appear to be non-cusp maximum at 88° meaning any difference in energy will be small.

The experimental torsional barrier is controversial - energies as low as as 250 kJ/-mol[114] and as high as 400 kJ/mol[180] have been suggested. In any case - as mentioned in Chapter 3 computational treatment of the torsion barrier likely needs to consider

relaxation of the C-C bond length coupled to dihedral angle rotation,[115] so even exact or near-exact benchmark calculations of this type would be expected to overestimate the experimentally observed value.

Fortunately, variational and diffusion Monte Carlo calculations[179] provide access to highly accurate theoretical estimates of the torsional barrier height, as reported in Table 7.4.

In agreement with previous literature reports, the single-reference RHF and CCSD(T) models do not accurately predict barrier heights, due to their inability to capture static correlation effects.[24,66] The statically-correlated SFS-NOCID and CASSCF(2,2) models already yield substantially more accurate predictions, but applying subsequent perturbation theory corrections brings them into very close agreement with the computational reference values. However, SFS-NOCID-PT2 calculations can be carried out for a fraction of the computational cost associated with MCQDPT2.

It has previously been shown that CCSD(T) calculations cannot adequately model static correlation effects near near the barrier[24,66] and consequently overestimate the barrier height. Therefore it isn't surprising that all multi-reference models yield lower torsional barriers than CCSD(T).

7.4.2 N → T Excitation Energy

Excitation energies are computed as the difference between the excited and ground state energies of the system at its equilibrium geometry:

$$E_{\text{exc}} = E(T/Z) - E(N) \tag{7.7}$$

When computed using statically correlated wavefunctions, this quantity is denoted $E_{\text{exc}}^{\text{stat}}$. Dynamic contributions to excitation energies are defined as:

$$E_{\text{exc}}^{\text{dyn}} = E_{\text{exc}} - E_{\text{exc}}^{\text{stat}} \tag{7.8}$$

Method	$E_{\text{exc}}^{\text{stat}}$	$E_{\text{exc}}^{\text{dyn}}$	E_{exc}
SFS-NOCISD/+PT2	410.3	-49.7	369.0
CASSCF*/MCQDPT2*	411.9	66.3	478.2
EOM-CCSD	–	–	435.4
icCAS-CI[157]	–	–	421.6
Experiment[160]			415

Table 7.5: Calculated, computational and experimental reference excitation energies for the N $\rightarrow$ T transition of ethylene. All calculations performed in this work use the cc-pVTZ basis set. The icCAS-CI wavefunction comes from a CI calculation performed in a complete active space of 12 electrons and 17 MOs in an augmented cc-pVDZ basis set basis set.[157,181] All energies in kJ/mol.

As Table 7.5, shows SFS-NOCISD-PT2 underestimates the N $\rightarrow$ T excitation energies by over 10%.

In both SFS-NOCISD-PT2 and MCQDPT, applying the second-order perturbation theory "correction" actually yields excitation energies further from the experimental and icCAS-CI reference values than the underlying statically correlated SFS-NOCISD and CASSCF(2,2) models.

Unlike with MCQDPT, applying perturbative corrections to SFS-NOCISD actually decreased the excitation energy, preferentially stabilizing the excited state over the ground state. This result is counterintuitive, because the electron density in the ground state is expected to be more compact, *i.e.* the electrons are closer together and therefore the effects of dynamic correlation are expected to be greater. This implies a problem with our implementation of the PT corrections to the SFS-NOCISD wavefunction. Likely this lies in the assumption that the perturbation does not need to be applied to off diagonal elements of the SFS-NOCISD Hamiltonian.

The diagonal block of the SFS-NOCISD Hamiltonian that contains the determinants, that form the T state includes off diagonal elements with a magnitude approximately 15% that of the diagonal elements (data not shown). This means that our assumption that the corrections to the off diagonal elements can be ignored is likely not valid for this state. In this case, the large off diagonal elements of the

Method	$E_{\text{exc}}^{\text{stat}}$	$E_{\text{exc}}^{\text{dyn}}$	E_{exc}
SFS-NOCISD/+PT2	1494.0	-92.3	1401.8
CASSCF*/MCQDPT2*	1438.6	-158.9	1279.8
EOM-CCSD	–	–	1257.9
EOM-CCSD(T)[182]	–	–	815.3
Experimental[165]	798		

Table 7.6: Calculated, computational and experimental reference excitation energies for the N $\rightarrow$ Z transition of ethylene. All energies in kJ/mol.

SFS-NOCISD Hamiltonian correspond to two configurations constructed from the same set of reference orbitals, which are hence orthogonal. This means an MP2 correction could be applied to these off-diagonal elements with any significant modifications to our algorithm. However, this approach could easily cause an unbalanced treatment of states in which states composed of primarily orthogonal orbitals are modelled better than those which are not.

7.4.3 N $\rightarrow$ Z Excitation Energy

Table 7.6 shows triply excited contributions to the wavefunction must be accounted for to accurately compute the energy of the Z state and therefore obtain accurate transition energies. All methods that include only double excitations (or second order perturbative corrections) overestimate the excitation energy by more than 50%.

SFS-NOCISD overestimates the energy relative to CASSCF*(2,2) and EOM-CCSD by 10% again. Unlike in the case of the T state, the perturbation theory correction at least shifts the computed excitation energies towards the experimental and computational reference values.

7.5 Conclusion

In this chapter, we have shown that simple second-order perturbative corrections to SF-NOCI models yield energies which are very close to those obtained from the more

sophisticated MCQDPT approach. For ground state dissociation energies (LiH) and torsional barrier heights (ethylene) both the SF-NOCI-PT2 and MCQDPT2 models closely reproduced experimental and literature reference values.

Similarly, for transition energies to the lowest-lying triplet and singlet excited states at equilibrium, all models that account for both dynamic and static correlation reproduce benchmark computational values to within 30 kJ/mol or ~ 0.3 eV. For the experimentally observable $X^1\Sigma^+ \rightarrow A^1\Sigma^+$ transition, all "fully correlated" predictions lie within 10 kJ/mol or ~ 0.1 eV of the experimentally-determined value.

However, for the excited states of ethylene, the SFS-NOCI-PT2 energies often differed significantly from the MCQDPT2 ones and both approaches either diverged from the experimental value (N $\rightarrow$ T transition) or did not substantially improve the statically correlated result (N $\rightarrow$ T transition). In these cases, further work is required to understand where the core assumptions underpinning this model fail, and whether model limitations are best addressed by varying the spin-flip molecular orbital basis, or allowed excitations within those orbitals, or whether off-diagonal MP2 corrections are required, or whether "diagonalise-and-then-perturb" models may be more appropriate.

Chapter 8

Conclusions and Further Work

8.1 Summary

8.1.1 SF-NOCI: A Minimal-Determinant Static Correlation Model

In this thesis, we set out to design and implement a novel minimal-determinant electronic structure model that contains only the electronic configurations / determinants required to capture static electron correlation effects, and provide a solid foundation for subsequent dynamic correlation corrections to be applied.

In physical terms, this means excluding determinants that would otherwise correct for the mean-field approximation (capture dynamic correlation effects) or poor quality reference orbitals (the unoptimized HF virtuals that are a consequence of HF being a single-reference method). The resultant wavefunctions should correspond more closely to chemists' intuitions about molecular electronic structure, thereby allowing for easier and more chemically meaningful analysis.

In Chapter 1, we identify 3 key requirements:

1. The configuration expansion must be concise and simple to specify - this restrains the scaling of the method and helps ensure there is a clear chemical

interpretation to the result.

2. Closely related to 1; the wavefunction must make use of optimized orbitals - this is required to allow a concise wavefunction to be accurate enough to be useful.

3. Finally, the optimization of the MOs and configuration series must be numerically stable to make the method practical without extensive trial and error.

No existing methods satisfy all three criteria. While general multi-configurational SCF (MCSCF) methods provide access to high quality optimized orbitals within concise configuration expansions, setting up these calculations requires extensive trial and error and problem-specific knowledge. Complete active space SCF (CASSCF) methods partly solve the configuration selection problem by allowing all possible excitations of electrons contained within a pre-defined active space of occupied and virtual orbitals. However, they also require careful problem-specific selection of orbitals for inclusion in the active space, and the resultant configuration expansion generally contains more determinants than strictly required to capture *only* static correlation.

We identified the non-orthogonal spin-flip configuration interaction method developed by the Head-Gordon group[75] as a suitable foundation for building such a method. High-spin reference calculations are used to obtain semi-optimized MOs for excited determinants with the same molecular orbital occupancy, avoiding the need to use unoptimized virtual orbitals from single-reference HF calculations. In previous work, this has been used to accelerate the convergence of "conventional" CI expansions.[64]

In Chapter 2, three schemes for selecting minimal sets of configurations to include in the NOCI expansion are proposed, along with recommendations on the realm of applicability of each approach. All three schemes are designed to satisfy the requirements for an ideal minimal-determinant static correlation model listed above.

8.1.2 Dynamic Correlation: Second Order Perturbation Theory

In contrast to methods like SF-NOCI that use a minimal set of carefully selected configurations with optimized orbitals to capture static correlation, dynamic correlation models typically comprise many configurations each making relatively small contributions to describing the subtle ways that electrons explicitly avoid one another, beyond the mean-field approximation.

Second-order Möller-Plesset perturbation theory provides an economical way of accounting for these contributions. In Chapter 6 and Appendix B, we describe how second order perturbation theory may be adapted to apply to SF-NOCI wavefunctions. This new class of methods is referred to as SF-NOCI-PT2.

Following literature precedent,[174] perturbative corrections are applied only to the diagonal elements of the NOCI matrix. The use of state-specific optimized orbitals makes the NOCI matrix even more diagonal-dominant than corresponding CI matrices, and this is shown to be a very effective approximation in the SF-NOCI context.

8.1.3 Model Systems: LiH and Ethylene

SF-NOCI and SF-NOCI-PT2 models were applied to two model systems – dissociating LiH and twisting ethylene – both of which are known to require static correlation contributions to ensure qualitatively correct wavefunctions at all points along the bond dissociation (LiH)[149,154] and C–C bond rotation (ethylene)[112,113] coordinates. We also assessed the ability of these models to predict vertical transition energies to electronically-excited states at the equilibrium geometry for each molecule.

For recovering static correlation energies and determining wavefunction compositions, our SF-NOCI models performed better than SFCI and similar to CASSCF wavefunctions of equivalent size. Including dynamic correlation corrections using SF-NOCI-PT2 often produced results close to both experimental values and high-level reference calculations, particularly for ground state quantities.

Our singly-spin-flipped minimal-determinant approach is not as well suited to describing excited states. The V state of ethylene, which is well known to mix strongly with higher energy configurations,[156,161] was particularly problematic.

Clearly, methods that select configurations pertinent to ground state dissociation and twisting processes are not necessarily well suited to describing excited states, particularly if the excited states included in the SF-NOCI basis themselves lie close in energy to even higher excited states that are not included in the SF-NOCI expansion. However, the very simple FR-SFS-NOCIS method was sufficient to predict vertical transition energies to the lowest-lying triplet and singlet states of LiH, illustrating the potential of this approach.

8.2 Further Work

Although SF-NOCI-PT2 accurately models ground state potential curves and some excited state transition energies of our two model systems, there is substantial work to be done before we could recommend it as a general purpose ground and excited state method. We applied our methods to only two small molecules each with two "active" electrons in the HOMO; so only a single level of spin-flip was required to obtain semi-optimized HOMO and LUMO orbitals to form the excited states in the SF-NOCI determinant basis. There is therefore a need to explore how our method applies to a wider range of systems, especially in situations such as multiple bond breaking, which will require higher levels of spin flipping and excitation.

The failure of SFS-NOCISD to model the V state of ethylene also raises the question of how to select minimal orbital spaces for excited state calculations. This is a question that remains unsolved after a long history in the MCSCF literature and SF-NOCI does not inherently make the problem any easier. However, the relatively cheap cost of including determinants in the SF-NOCI expansion and the ease with which higher energy reference orbitals can be found, make it a useful tool in exploring the nature of excited states involving multiple higher energy configurations.

Additionally SF-NOCI enables a more straightforward exploration of the space of possible CI expansions by removing one of the additional degrees of freedom in MCSCF calculations, namely the choice of configuration weighting during reference orbital optimization.

Finally, there is the issue of how the MP2 correction is implemented. The case of the T excited state showed that in some situations our dynamic correlation correction can give energies which are qualitatively different to more rigorous multi-configuration perturbation theories. We suggested that this error arises from large off diagonal elements of the SF-NOCI Hamiltonian however at present we have no real detailed understanding of how our approximation compares with a full treatment of SF-NOCI and how its limitations can be ameliorated. One possibility mentioned in Chapter 7 is to apply the perturbative correction to elements of the CI matrix that share a common set of reference orbitals. Additionally our approximation should be compared to the full NOCI matrix approach to NOCI-PT2 developed by the Head-Gordon group,[183] to develop a more complete understand of its trade-offs.

8.3 Conclusions

Over the course of this thesis, we have developed and demonstrated a method that shows promise in achieving the goals we set out in Chapter 1. SF-NOCI models yield qualitatively correct wavefunctions in many of the cases we have explored. Furthermore, simple dynamic correlation corrections often yield relative energies and transition energies comparable in accuracy to those obtained using more expensive methods. However, this was not true in all cases. For excited states of ethylene, SF-NOCI-PT2 transition energies differed substantially from experimental data and benchmark calculations. More work to determine the limitations and applicability of our assumptions is needed in these cases.

Nonetheless, this work demonstrates the value of our fundamental philosophy - clearly separating static and dynamic correlation contributions to the overall wave-

function and energy, each handled using methods tailored to provide a minimal treatment of each form of interaction. Initial applications show the potential of this approach to provide simple and accurate models for challenging chemical systems. With our SF-NOCI-PT2 implementation, we have produced a tool that is capable of modelling multi-reference systems more cheaply and conceptually simply than existing methods of comparable accuracy.

Appendix A

SF-NOCI Code

This appendix contains an implementation of the NOCI method taken from our pychem software developed for this thesis. It has direct dependencies on only two external libraries - the numpy[184] and scipy[185] packages - and a number of other modules within the pychem program.

- structures: Data structures to represent the molecules and electronic states, as well as calculation settings.
- hartree_fock: Functions for calculating the matrices associated with the HF method using the objects provided in structures.
- mp2: An implementation of the standard MP2 method for electronic states, as described in chapter 6.
- printf: Utilities for printing output in a structured way.

The full pychem program can be found at: github.com/dlc62/pychem.

```
# System libraries
import numpy as np
from scipy.linalg import eigh as gen_eig
from collections import namedtuple
from copy import copy
# Custom code
from Util import printf, structures
```

```
from Util.structures import Spin
import hartree_fock as hf
import mp2

# Magic number defining what values is conisdered a zero overlap
# for the purposes of MO orthogonality
NOCI_thresh = 1e-8

#-------------------------------------------------------------#
#                Set up required structures                   #
#-------------------------------------------------------------#

ZeroOverlap = namedtuple("Zero", "index, spin")

class CoDensityState:
   # # A state-like object used for calculating and storing the pseudo Fock matrix
    def __init__(self, n_orbitals, alpha_density, beta_density):
        self.Alpha = Matrices(n_orbitals)
        self.Beta = Matrices(n_orbitals)
        self.Total = Matrices(n_orbitals,total=True)
        self.Alpha.Density = alpha_density
        self.Beta.Density = beta_density
        self.Total.Density = alpha_density + beta_density

class Matrices:
    def __init__(self,n_orbitals,total=False):
        self.Density = np.zeros((n_orbitals,) * 2)
        if not total:
            self.Exchange = np.zeros((n_orbitals,) * 2)
        else:
            self.Coulomb = np.zeros((n_orbitals,) * 2)

#-------------------------------------------------------------#
#                       Main routine                          #
#-------------------------------------------------------------#

def pprint_spin_flip_states(states):
    string = "\n"
    for state in states:
        string += str(state) + "\n"
    return string

def do(settings, molecule):

    if "SF" in molecule.ExcitationType:
        make_natural_orbitals(molecule)
        CI_states = [make_SF_NOCI_state(state, molecule.States)
                     for state in molecule.SpinFlipStates]
    else:
        CI_states = molecule.States

    dims = len(CI_states)              # Dimensionality of the CI space
    CI_matrix = np.zeros((dims, dims))
    CI_overlap = np.zeros((dims, dims))
    nA = molecule.NAlphaElectrons; nB = molecule.NBetaElectrons

    # Building the CI matrix
    for i, state1 in enumerate(CI_states):
        for j, state2 in enumerate(CI_states[:i+1]):

            alpha, bover = biorthogonalize(state1.Alpha.MOs, state2.Alpha.MOs, molecule.Overlap, nA)
            beta, aover = biorthogonalize(state1.Beta.MOs, state2.Beta.MOs, molecule.Overlap, nB)

            # Calculate the core fock matrix for the state transformed into the MO basis
            alpha_core = alpha[0].T.dot(molecule.Core).dot(alpha[1])
            beta_core = beta[0].T.dot(molecule.Core).dot(beta[1])

            # Get the diagonal elements and use them to calculated the overlap, reduced overlap
            # and the list of zero overlaps
            alpha_overlaps = np.diagonal(alpha[0].T.dot(molecule.Overlap).dot(alpha[1]))
            beta_overlaps = np.diagonal(beta[0].T.dot(molecule.Overlap).dot(beta[1]))
            state_overlap = np.product(alpha_overlaps) * np.product(beta_overlaps) * aover * bover
```

```
            reduced_overlap, zeros_list = process_overlaps(1, [], alpha_overlaps, Spin.Alpha)
            reduced_overlap, zeros_list = process_overlaps(reduced_overlap, zeros_list, beta_overlaps, Spin.Beta)
            reduced_overlap *= aover * bover

            num_zeros = len(zeros_list)

            # Calculate the Hamiltonian matrix element for this pair of states
            # And find the combined state
            if num_zeros == 0:
                elem,state = no_zeros(molecule, alpha, beta, alpha_overlaps, beta_overlaps,
                                      alpha_core, beta_core)
            elif num_zeros == 1:
                elem, state = one_zero(molecule, alpha, beta, alpha_overlaps, beta_overlaps,
                                       alpha_core, beta_core, zeros_list[0])
            elif num_zeros == 2:
                elem, state = two_zeros(molecule, alpha, beta, zeros_list)
            else:  # num_zeros > 2
                elem = 0

            elem *= reduced_overlap
            elem += molecule.NuclearRepulsion * state_overlap
            CI_matrix[i,j] = CI_matrix[j,i] = elem
            CI_overlap[i,j] = CI_overlap[j,i] = state_overlap

            # Only apply the mp2 correction to the diagonal elements
            if i == j and settings.Method == "NOCI-PT2":
                CI_matrix[i,j] += get_mp2_corrections([state], molecule, settings)[0]

    # Print the Hamiltonian and State overlaps before attemption to solve the eigenvalue problem
    # so we still get information if it fails
    printf.delimited_text(settings.OutFile," NOCI output ")
    printf.text_value(settings.OutFile, " Hamiltonian ", CI_matrix, " State overlaps ", CI_overlap)

    # Solve the generalized eigenvalue problem
    energies, wavefunctions = gen_eig(CI_matrix, CI_overlap)

    molecule.NOCIEnergies = energies
    molecule.NOCIWavefunction = wavefunctions

    printf.text_value(settings.OutFile, " States ", wavefunctions, " NOCI Energies ", energies)

#-------------------------------------------------------------------#
#     Functions for rearranging the HF orbitals as required         #
#-------------------------------------------------------------------#
def assemble_orbitals(occupancies, optimized):
    new_MOs = np.empty(optimized.MOs.shape)
    new_energies = np.empty(np.shape(optimized.Energies))

    optimized_count = 0
    unoptimized_count = sum(occupancies)
    for i, occ in enumerate(occupancies):
        if occ == 0:
            new_MOs[:,unoptimized_count] = optimized.MOs[:,i]
            new_energies[unoptimized_count] = optimized.Energies[i]
            unoptimized_count += 1
        elif occ == 1:
            new_MOs[:,optimized_count] = optimized.MOs[:,i]
            new_energies[optimized_count] = optimized.Energies[i]
            optimized_count += 1
        else:
            raise ValueError("Trying to construct a NOCI state without a suitable HF optimized orbital")

    return new_MOs, new_energies

def optimize_active_space(AO_overlaps, reference, nBeta, NOs):
    # Use the doubly occupied space of the ground state as the reference orbitals
    overlap_operator = AO_overlaps.dot(reference[:,:nBeta])
    B = NOs.T.dot(overlap_operator).dot(overlap_operator.T).dot(NOs)
    _lambda, coeffs = np.linalg.eigh(B)
    new_NOs = NOs.dot(coeffs.T)

    # Reverse the NOs to get them in order of assending occupancy
    new_NOs = new_NOs[:,::-1]
```

```
    return new_NOs

def find_UHF_natural_orbitals(state, S):
    # See J. Chem. Phys. 1988, 88(8), 4926
    half_density_matrix = S.dot(state.Total.Density/2).dot(S)
    NO_vals, NO_vects = np.linalg.eigh(half_density_matrix)
    NO_coeffs = np.linalg.inv(S).dot(NO_vects)

    # Get the NOs in ascending order of occupancy
    NO_coeffs = NO_coeffs[:,::-1]
    NO_vals = NO_vals[::-1]

    return NO_coeffs, NO_vals

def make_natural_orbitals(molecule):
    reference_orbitals = None
    for i, state in enumerate(molecule.States):
        NOs, occ = find_UHF_natural_orbitals(state, molecule.S)
        # Save the ground state NOs as a reference
        if i == 0:
            reference_orbitals = NOs

        # Now we need to account for degeneracy in the singly occupied space
        # If there is more than one singly occupied orbital we need to optimize
        # their overlaps with the reference orbitals
        # Find the singly occupied orbitals
        idx = [i for i,x in enumerate(occ) if np.isclose(x, 0.5)]
        if len(idx) > 1:
            NOs[:,idx] = optimize_active_space(
                molecule.Overlap,
                reference_orbitals,
                molecule.NBetaElectrons,
                NOs[:,idx])

        state.Alpha.MOs, state.Beta.MOs = copy(NOs)

def make_SF_NOCI_state(spin_flip_state, hf_states):
    # Select the required HF state
    NOrbs = len(spin_flip_state[1])
    HF_state = hf_states[spin_flip_state[0]]

    new_alpha, new_alpha_energies = assemble_orbitals(spin_flip_state[1], HF_state.Alpha)
    new_beta, new_beta_energies = assemble_orbitals(spin_flip_state[2], HF_state.Alpha)

    new_state = structures.ElectronicState(spin_flip_state[1], spin_flip_state[2], NOrbs)
    new_state.Alpha.MOs = new_alpha
    new_state.Alpha.Energies = new_alpha_energies
    new_state.Beta.MOs = new_beta
    new_state.Beta.Energies = new_beta_energies
    new_state.TotalEnergy = HF_state.TotalEnergy
    new_state.Energy = HF_state.Energy

    return new_state

#-----------------------------------------------------------------------#
#  Functions for computing overlaps, densities, biorthogonalized MOs  #
#-----------------------------------------------------------------------#
def biorthogonalize(old_MOs1, old_MOs2, AO_overlaps, nElec):
    # This function finds the Lowdin Paired Orbitals for two sets of MO coefficents
    # using a singular value decomposition, as in J. Chem. Phys. 140, 114103
    # Note this only returns the MO coeffs corresponding to the occupied MOs """

    # Finding the overlap of the occupied MOs
    MOs1 = copy(old_MOs1[:,:nElec])
    MOs2 = copy(old_MOs2[:,:nElec])
    MO_overlaps = MOs1.T.dot(AO_overlaps).dot(MOs2)

    # Check if the orbitals are already paried
    if is_biorthogonal(MOs1, MOs2, AO_overlaps):
        return [MOs1, MOs2], 1

    U, _sigma, Vt = np.linalg.svd(MO_overlaps)
```

```
    over = np.linalg.det(U) * np.linalg.det(Vt)

    # Transforming each of the determinants into a biorthogonal basis
    new_MOs1 = MOs1.dot(U)
    new_MOs2 = MOs2.dot(Vt.T)

    assert is_biorthogonal(new_MOs1, new_MOs2, AO_overlaps)

    return [new_MOs1, new_MOs2], over

def is_biorthogonal(MOs1, MOs2, AO_overlaps):
    size = MOs1.shape[1]
    MO_overlaps = MOs1.T.dot(AO_overlaps).dot(MOs2)
    residuals = np.abs(MO_overlaps) - np.eye(size, size)
    biorthogonal = residuals.max() < NOCI_thresh

    return biorthogonal

def make_weighted_density(MOs, overlaps):
    nOrbs = np.shape(MOs)[1]
    density = np.zeros((nOrbs, nOrbs))
    for i, overlap in enumerate(overlaps):
        if overlap > NOCI_thresh:
            P = np.outer(MOs[0][:,i], MOs[1][:,i])
            density += P / overlap
    return density

def process_overlaps(reduced_overlap, zeros_list, overlaps, spin):
    # Builds up the list of zero values as a list of (zero, spin) tuples
    # as well as the reduced overlap value
    for i, overlap in enumerate(overlaps):
        if np.abs(overlap) > NOCI_thresh:
            reduced_overlap *= overlap
        else:
            zeros_list.append(ZeroOverlap(i, spin))
    return reduced_overlap, zeros_list

def resize_array(src, dest, fill=0):
    """ Makes array src the same size as array dest, the old array is embeded in
     the upper right corner and the other elements are zero. Only works for
     for projecting a vector into a vector or a matrix into a matrix """
    old_shape = np.shape(src)
    new_array = np.full_like(dest, fill)
    if len(old_shape) == 2:              # Matrix
        height, width = old_shape
        new_array[:height, :width] = src
    else:                                # Vector
        length = old_shape[0]
        new_array[:length] = src
    return new_array

#-----------------------------------------------------------------#
#  Functions for calculating NOCI matrix elements, not including  #
#  reduced overlap term which is accounted for later              #
#-----------------------------------------------------------------#

def no_zeros(molecule, alpha, beta, alpha_overlaps, beta_overlaps, alpha_core, beta_core):

    W_alpha = make_weighted_density(alpha, alpha_overlaps)
    W_beta = make_weighted_density(beta, beta_overlaps)
    state = CoDensityState(molecule.NOrbitals, W_alpha, W_beta)
    hf.make_coulomb_exchange_matrices(molecule, state)

    elem = inner_product(W_alpha + W_beta, state.Total.Coulomb)
    elem += inner_product(W_alpha, state.Alpha.Exchange)
    elem += inner_product(W_beta, state.Beta.Exchange)
    elem *= 0.5

    # Add the one electron terms
    for i in range(molecule.NAlphaElectrons):
        if alpha_overlaps[i] > NOCI_thresh:
            elem += alpha_core[i,i] / alpha_overlaps[i]
```

```
    for i in range(molecule.NBetaElectrons):
        if beta_overlaps[i] > NOCI_thresh:
             elem += beta_core[i,i] / beta_overlaps[i]

    return elem, state

def one_zero(molecule, alpha, beta, alpha_overlaps, beta_overlaps, alpha_core, beta_core, zero):
    zero_index = zero.index

    # Making all the required Codensity matrices
    W_alpha = make_weighted_density(alpha, alpha_overlaps)
    W_beta = make_weighted_density(beta, beta_overlaps)
    P_alpha = np.outer(alpha[0][:,zero_index], alpha[1][:,zero_index])
    P_beta = np.outer(beta[0][:,zero_index], beta[1][:,zero_index])

    state = CoDensityState(molecule.NOrbitals, W_alpha, W_beta)
    hf.make_coulomb_exchange_matrices(molecule, state)

    active_exchange = state.Alpha.Exchange if zero.spin == Spin.Alpha else state.Beta.Exchange
    P_active = P_alpha if zero.spin == Spin.Alpha else P_beta
    active_core = alpha_core if zero.spin == Spin.Alpha else beta_core
    elem = inner_product(P_active, state.Total.Coulomb) + inner_product(P_active, active_exchange)
    elem += active_core[zero_index, zero_index]

    return elem, state

def two_zeros(molecule, alpha, beta, zeros_list):
    i, spin = zeros_list[0]

    P_alpha = np.outer(alpha[0][:,i], alpha[1][:,i])
    P_beta = np.outer(beta[0][:,i], beta[1][:,i])
    state = CoDensityState(molecule.NOrbitals, P_alpha, P_beta)
    hf.make_coulomb_exchange_matrices(molecule, state)
    active_exchange = state.Alpha.Exchange if spin == Spin.Alpha else state.Beta.Exchange
    active_P = P_alpha if spin == Spin.Alpha else P_beta

    elem = inner_product(active_P, state.Total.Coulomb) + inner_product(active_P, active_exchange)

    return elem, state

def inner_product(mat1, mat2):
    product = mat1.dot(mat2.T)
    return np.trace(product)

def get_mp2_corrections(hf_states, molecule, settings):
    mp2_corrections = []

    for i, mp2_state in enumerate(hf_states):
        hf.make_fock_matrices(molecule, mp2_state)
        hf.make_MOs(molecule, mp2_state)

        mp2_corrections.append(mp2.do(settings, molecule, [mp2_state]))

    return mp2_corrections
```

Appendix B

Non-Orthogonal MP2

B.1 Rayleigh-Schrödinger Perturbation Theory

All perturbative techniques are based upon a few simple assumptions. The first of these is that we have a mathematical expression for a physical quantity for which we are unable to obtain an exact solution. The next assumption is that this physical quantity may be broken down into a part which can be solved exactly and a troublesome part which has no analytic solution. This "perturbation" is assumed to be relatively small in comparison to the soluble portion of the problem. The true solution is expanded in a basis of functions obtained as analytic solutions to the solvable part of the problem.

Rayleigh-Schrödinger perturbation theory requires no *a priori* assumptions about the form of the Hamiltonian ($\hat{H}$), or its decomposition into an analytically solvable part ($\hat{H}_0$) and perturbative part $\hat{V}$, or the nature of the analytic solutions to $\hat{H}_0$ ($E_n^{(0)}$, $\Psi_n^{(0)}$). It is formulated completely generally, introducing only the perturbation parameter λ to keep track of orders of correction terms.

$$\hat{H} = \hat{H}_0 + \lambda\hat{V} \tag{B.1}$$

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \lambda^3 E_n^{(3)} + \ldots \tag{B.2}$$

$$\Psi_n = \Psi_n^{(0)} + \lambda \Psi_n^{(1)} + \lambda^2 \Psi_n^{(2)} + \lambda^3 \Psi_n^{(3)} + \ldots \tag{B.3}$$

Substituting these expressions into $\hat{H}\Psi_n = E_n\Psi_n$ and equating powers of lambda on both sides yields (to 2nd order):

$$\hat{H}_0\Psi_n^{(0)} = E_n^{(0)}\Psi_n^{(0)} \tag{B.4}$$

$$\hat{H}_0\Psi_n^{(1)} + \hat{V}\Psi_n^{(0)} = E_n^{(0)}\Psi_n^{(1)} + E_n^{(1)}\Psi_n^{(0)} \tag{B.5}$$

$$\hat{H}_0\Psi_n^{(2)} + \hat{V}\Psi_n^{(1)} = E_n^{(0)}\Psi_n^{(2)} + E_n^{(1)}\Psi_n^{(1)} + E_n^{(2)}\Psi_n^{(0)} \tag{B.6}$$

Left-multiplying by $\Psi_n^{(0)}$ and integrating both sides, assuming and/or imposing $\langle\Psi_n^{(0)}|\hat{H}_0|\Psi_n^{(m)}\rangle = 0$ and $\langle\Psi_n^{(0)}|\Psi_n^{(m)}\rangle = 0$ yields:

$$E_n^{(0)} = \langle\Psi_n^{(0)}|\hat{H}_0|\Psi_n^{(0)}\rangle \tag{B.7}$$

$$E_n^{(1)} = \langle\Psi_n^{(0)}|\hat{V}|\Psi_n^{(0)}\rangle \tag{B.8}$$

$$E_n^{(2)} = \langle\Psi_n^{(0)}|\hat{V}|\Psi_n^{(1)}\rangle \tag{B.9}$$

$$\vdots$$

$$E_n^{(m)} = \langle\Psi_n^{(0)}|\hat{V}|\Psi_n^{(m-1)}\rangle \tag{B.10}$$

If the $\Psi_n^{(0)}$ form a complete, orthonormal set, then they also form a basis for expanding higher order wavefunction correction terms, $\Psi_n^{(m)}$:

$$\Psi_n^{(m)} = \sum_l C_{n,l}^{(m)}\Psi_l^{(0)} \tag{B.11}$$

where

$$C_{n,l}^{(m)} = \langle\Psi_l^{(0)}|\Psi_n^{(m)}\rangle \tag{B.12}$$

For $m = 1$, the required coefficients may be obtained using only the 0th order solutions. Left-multiplying equation (B.5) by $\Psi_l^{(0)}$, inserting the resolution of the identity $|\Psi_l^{(0)}\rangle\langle\Psi_l^{(0)}|$ into the first term to the right of the Hamiltonian operator,

and integrating gives:

$$(E_n^{(0)} - E_l^{(0)}) \langle \Psi_l^{(0)} | \Psi_n^{(1)} \rangle = \langle \Psi_l^{(0)} | \hat{V} | \Psi_n^{(0)} \rangle \tag{B.13}$$

and so

$$C_{n,l}^{(1)} = \frac{\langle \Psi_l^{(0)} | \hat{V} | \Psi_n^{(0)} \rangle}{(E_n^{(0)} - E_l^{(0)})} \tag{B.14}$$

B.2 Single-reference MP2 derivation

Within an electronic structure context, the molecular Hamiltonian is partitioned into one- and two-electron terms, with a perturbation parameter, λ, labelling the perturbative two-electron part:

$$\hat{H} = \hat{H}_0 + \lambda \hat{V} \tag{B.15}$$

where $\hat{H}_0$ is the one-electron Fock operator:

$$\begin{aligned} \hat{H}_0 &= \sum_i \hat{f}(i) \\ &= \sum_i \hat{h}(i) + \hat{V}^{\mathrm{HF}}(i) \end{aligned} \tag{B.16}$$

where $\hat{h}$ represents the core Hamiltonian (kinetic energy and nucleus-electron potential terms) and $\hat{V}^{\mathrm{HF}}(i)$ represents the mean-field Coulomb and exchange potentials experienced by an electron in molecular orbital i. It does not account for Coulomb and exchange interactions between orbitals that are normally also included in the HF energy expression. These are captured – even within HF theory – through the two-electron term of the Hamiltonian, which has been separated out here for the purposes of developing perturbation theory:

$$\hat{V} = \sum_{i<j} r_{ij}^{-1} - \sum_i \hat{V}^{\mathrm{HF}}(i) \tag{B.17}$$

In this perturbation theory formulation, the Hartree-Fock energy is the sum of the 0th and 1st order perturbation theory terms, equations (B.7) and (B.8). Only the ground state, $n = 0$ is physically meaningful:

$$E_{\mathrm{HF}} = E_0^{(0)} + E_0^{(1)} \tag{B.18}$$

$$= \langle \Psi_0^{(0)} | \hat{H}_0 | \Psi_0^{(0)} \rangle + \langle \Psi_0^{(0)} | \hat{V} | \Psi_0^{(0)} \rangle \tag{B.19}$$

$$= \langle \Psi_0^{(0)} | \hat{H}_0 + \hat{V} | \Psi_0^{(0)} \rangle \tag{B.20}$$

Note that this is still a mean-field method despite the inclusion of the r_{12}^{-1} term in the Hamiltonian, because the Hartree-Fock wavefunctions are obtained within the mean-field approximation.

Perturbative corrections to the wavefunction – at any/all orders – are expanded in a basis of excited Hartree-Fock determinants, Φ, in a context-specific application of equation (B.11):

$$\Psi_0^{(m)} = \sum_{i,a} C_{i,a}^{(m)} \Phi_i^a + \sum_{i<j,a<b} C_{i,j,a,b}^{(m)} \Phi_{i,j}^{a,b} + \sum_{i<j<k,a<b<c} C_{i,j,k,a,b,c}^{(m)} \Phi_{i,j,k}^{a,b,c} + \dots \tag{B.21}$$

where $\Psi_0^{(0)} = \Phi$, i and j index occupied orbitals from which electrons originate and a,b index the virtual into which they are promoted. The required coefficients for the first-order correction to the wavefunction ($m = 1$) are computed according to equation (B.14), with $n = 0$ and $l = i, a$ (singly-excited determinants) or $l = i, j, a, b$ (doubly-excited determinants).

$$C_{i,a}^{(1)} = \frac{\langle \Phi_i^a | \hat{V} | \Phi \rangle}{E_0 - E_i^a} \tag{B.22}$$

$$C_{i,j,a,b}^{(1)} = \frac{\langle \Phi_{i,j}^{a,b} | \hat{V} | \Phi \rangle}{E_0 - E_{i,j}^{a,b}} \tag{B.23}$$

All singly-excited integrals, $\langle \Phi_i^a | \hat{V} | \Phi \rangle$, evaluate to 0 when Φ_i^a are orthonormal to Φ, by Brillouin's theorem. Therefore, $C_{i,a}^{(1)} = 0$. Similarly, all terms with three or more

indices evaluate to 0 because the MOs unaffected by the two electron operator $\hat{V}$ are orthogonal. This leaves:

$$\Psi_0^{(1)} = \sum_{i<j,a<b} \Phi_{i,j}^{a,b} \frac{\langle \Phi_{i,j}^{a,b} | \hat{V} | \Phi \rangle}{E_0 - E_{i,j}^{a,b}} \tag{B.24}$$

Substituting (B.24) into (B.9), recalling that $\Psi_0^{(0)} = \Phi$, gives the MP2 energy expression:

$$E_0^{(2)} = \sum_{i,j,a,b} \langle \Phi | \hat{V} | \Phi_{i,j}^{a,b} \rangle \frac{\langle \Phi_{i,j}^{a,b} | \hat{V} | \Phi \rangle}{E_0 - E_{i,j}^{a,b}} \tag{B.25}$$

$$= \sum_{i<j,a<b} \frac{| \langle \Phi | \hat{V} | \Phi_{i,j}^{a,b} \rangle |^2}{E_0 - E_{i,j}^{a,b}} \tag{B.26}$$

Recalling that we have partitioned the Hamiltonian such that $\hat{H}_0$ is the one-electron Fock operator, the energy denominator is simply computed as a difference in molecular orbital energies (eigenvalues of Fock matrix) associated with moving electrons from occupied orbitals i and j to unoccupied/virtual orbitals a and b:

$$E_0 - E_{i,j}^{a,b} = -((\epsilon_a + \epsilon_b) - (\epsilon_i + \epsilon_j)) \tag{B.27}$$

$$= \epsilon_i + \epsilon_j - \epsilon_a + \epsilon_b \tag{B.28}$$

This term will clearly be negative (stabilizing) because the energies of orbitals a and b are higher than those of i and j.

Finally, it remains to evaluate the integrals in the numerator:

$$\langle \Phi | \hat{V} | \Phi_{i,j}^{a,b} \rangle = \langle \Phi | r_{12}^{-1} | \Phi_{i,j}^{a,b} \rangle - \langle \Phi | \hat{V}^{\mathrm{HF}} | \Phi_{i,j}^{a,b} \rangle \tag{B.29}$$

All integrals over the one-electron $\hat{V}^{\mathrm{HF}}$ operator evaluate to zero, so it only remains to evaluate the two-electron terms. This is achieved by expanding the wavefunction

using the Slater-Condon rules. Results are expressed here in chemists' notation:

$$\langle \Phi | \hat{V} | \Phi_{i,j}^{a,b} \rangle = [ia||jb] \tag{B.30}$$

Substituting (B.30) and (B.28) into (6.11) yields the familiar MP2 energy expression:

$$E_0^{(2)} = \sum_{i<j,a<b} \frac{|[ia||jb]|^2}{\epsilon_i + \epsilon_j - \epsilon_a + \epsilon_b} \tag{B.31}$$

$$= \frac{1}{4} \sum_{i,j,a,b} \frac{|[ia||jb]|^2}{\epsilon_i + \epsilon_j - \epsilon_a + \epsilon_b} \tag{B.32}$$

Finally, to evaluate these molecular Coulomb integrals, it is necessary to expand them in an AO basis. Naïvely, this is quite straightforward:

$$[ia||jb] = \sum_{\mu} \sum_{\nu} \sum_{\lambda} \sum_{\sigma} C_{i\mu} C_{a\nu} C_{j\lambda} C_{b\sigma} [\mu\nu||\lambda\sigma] \tag{B.33}$$

where $\mathbf{C}_i$ represents the vector of atomic orbital coefficients for orbital i, the sum runs over all atomic orbitals that contribute to each molecular orbital, and $[\mu\nu||\lambda\sigma]$ is a two electron repulsion integral evaluated over atomic orbital basis functions μ, ν, λ and σ.

In practice, the computational scaling of this procedure is prohibitive. This equation can be factorised into two half-transforms, as initially formulated by Head-Gordon *et al.*[186] and more recently by Baker and Wolinski[187] in the context of dealing with more general sets of molecular orbitals, including those obtained using unrestricted Hartree-Fock. This reduces the computational scaling of the procedure from $O(N_{\text{orbs}}^8)$ to $O(N_{\text{orbs}}^5)$.

For completeness, the key result from Baker and Wolinski is reproduced here:

$$[\mu a||\lambda b] = \mathbf{Y}_{ab}^{\mu\lambda} = \sum_{\nu,\sigma} \mathbf{C}_{\nu a}^{\mathrm{T}} \mathbf{X}_{\nu\sigma}^{\mu\lambda} \mathbf{C}_{\sigma b} = \mathbf{C}_{a}^{\mathrm{T}} \mathbf{X}^{\mu\lambda} \mathbf{C}_{b} \tag{B.34}$$

$$[ia||jb] = \mathbf{Z}_{ij}^{ab} = \sum_{\mu,\lambda} \mathbf{C}_{\mu i}^{\mathrm{T}} \mathbf{Y}_{\mu\lambda}^{ab} \mathbf{C}_{\lambda j} = \mathbf{C}_{i}^{\mathrm{T}} \mathbf{Y}^{ab} \mathbf{C}_{j} \tag{B.35}$$

where μ, ν, λ and σ index atomic orbitals, i and j index occupied molecular orbitals, a and b index virtual molecular orbitals, and $\mathbf{X}$ is the Coulomb-exchange matrix evaluated in an AO basis and $\mathbf{C}$ are MO coefficient vectors.

B.3 Multi-reference perturbation theories

There are three major classes of multi-reference perturbation theories, but common to all of them is a CASSCF reference wavefunction, defined by a set of CASSCF-optimized orbitals, a complete set of ground and excited state determinants within a pre-defined molecular orbital active space and their respective interaction energies (stored within the CI matrix), and the corresponding CI coefficients that define the wavefunction/s (obtained by diagonalizing the CI matrix). The MO and CI coefficients are obtained concurrently and iteratively.[188]

When formulating perturbation theory corrections, the overall space of excited determinants/wavefunctions is partitioned into a "reference space" in which configurational mixing is explicitly accounted for and an "outer space" of additional determinants/wavefunctions that contribute to perturbation theory wavefunction expansions. However, different multi-reference perturbation theory methods differ in how the perturbation theory corrections are applied:

- to individual determinants:
  - prior to diagonalization of CI matrix ("perturb then diagonalize"): quasi-degenerate perturbation theories (MRMP[189] and MCQDPT[190])

  - following diagonalization of CI matrix ("diagonalize then perturb"): state-specific multi-reference perturbation theories (CASPT2[191])

- to the CASSCF/CAS-CI wavefunction as a whole: n-electron valence perturbation theory (NEVPT[192,193])

All multi-reference perturbation theories in which corrections are applied to individual determinants suffer from the so-called "intruder state problem" in which approximate perturbative corrections to the CI basis states introduce accidental near-degeneracies, which is reflected in "unphysical" mixing within CI wavefunctions.

NEVPT overcomes this problem by using excited state CI wavefunctions rather than excited state determinants within the perturbation theory equations. As expected, this induces substantial additional algebraic and computational complexity, making this approach both harder to implement and more computationally intensive to execute.

Overall, multi-reference perturbation theories are non-trivial to both define and implement, even to second order.

B.4 SF-NOCI perturbation theory

Fortunately, our SF-NOCI implementation is more akin to HF-based single reference models than CASSCF-based multi-reference models. The NOCI expansion is comprised exclusively of HF-like basis states, rather than excited state electronic configurations. Therefore, the complexity and ambiguity of multi-reference MP2 models can be avoided. NEVPT simply doesn't apply, as there is no CASSCF wavefunction and no active space to make excitations to/from to form CASSCF excited states. This leaves only the option of applying MP2 corrections to individual CI matrix elements, either before or after diagonalization. Because we are working in an minimal determinant basis, there should be no opportunity for intruder states to arise.

Computing MP2 corrections for the diagonal elements of the NOCI matrix is trivial; these are simply found by applying the standard single-reference model to each HF basis state.

Deriving and computing MP2 corrections for the off-diagonal terms is harder. It requires a generalised form of RSPT for non-orthogonal wavefunctions, which will be developed using the integrated form of the Schrödinger equation rather than the differential form more commonly employed:

$$\langle \Psi_n | \hat{H} | \tilde{\Psi}_n \rangle = E_n \langle \Psi_n | \tilde{\Psi}_n \rangle \tag{B.36}$$

where Ψ and $\tilde{\Psi}$ are generally not orthogonal to one another.

The Hamiltonian is partitioned into a core one-electron term, $\hat{H}_0$, and a two-electron perturbation $\hat{V}$. Wavefunctions and energies are expanded as power series in the

perturbation parameter, λ:

$$\hat{H} = \hat{H}_0 + \lambda\hat{V} \tag{B.37}$$

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots \tag{B.38}$$

$$\Psi_n = \Psi_n^{(0)} + \lambda\Psi_n^{(1)} + \lambda^2\Psi_n^{(2)} + \ldots \tag{B.39}$$

$$\tilde{\Psi}_n = \tilde{\Psi}_n^{(0)} + \lambda\tilde{\Psi}_n^{(1)} + \lambda^2\tilde{\Psi}_n^{(2)} + \ldots \tag{B.40}$$

Substituting these expressions into (B.36) and equating terms at the same order of perturbation on both sides yields (to second order):

$$\langle\Psi_n^{(0)}|\hat{H}_0|\tilde{\Psi}_n^{(0)}\rangle = E_n^{(0)}\,\langle\Psi_n^{(0)}|\tilde{\Psi}_n^{(0)}\rangle \tag{B.41}$$

$$\begin{aligned}\langle\Psi_n^{(0)}|\hat{H}_0|\tilde{\Psi}_n^{(1)}\rangle + \langle\Psi_n^{(1)}|\hat{H}_0|\tilde{\Psi}_n^{(0)}\rangle + = E_n^{(0)}(\langle\Psi_n^{(0)}|\tilde{\Psi}_n^{(1)}\rangle + \langle\Psi_n^{(1)}|\tilde{\Psi}_n^{(0)}\rangle) + \\ \langle\Psi_n^{(0)}|\hat{V}|\tilde{\Psi}_n^{(0)}\rangle \qquad E_n^{(1)}\,\langle\Psi_n^{(0)}|\tilde{\Psi}_n^{(0)}\rangle\end{aligned} \tag{B.42}$$

$$\begin{aligned}\langle\Psi_n^{(0)}|\hat{H}_0|\tilde{\Psi}_n^{(2)}\rangle + \langle\Psi_n^{(1)}|\hat{H}_0|\tilde{\Psi}_n^{(1)}\rangle + \langle\Psi_n^{(2)}|\hat{H}_0|\tilde{\Psi}_n^{(0)}\rangle + = E_n^{(0)}(\langle\Psi_n^{(0)}|\tilde{\Psi}_n^{(2)}\rangle + \langle\Psi_n^{(1)}|\tilde{\Psi}_n^{(1)}\rangle + \langle\Psi_n^{(2)}|\tilde{\Psi}_n^{(0)}\rangle) + \\ \langle\Psi_n^{(0)}|\hat{V}|\tilde{\Psi}_n^{(1)}\rangle + \langle\Psi_n^{(1)}|\hat{V}|\tilde{\Psi}_n^{(0)}\rangle \qquad E_n^{(1)}(\langle\Psi_n^{(0)}|\tilde{\Psi}_n^{(1)}\rangle + \langle\Psi_n^{(1)}|\tilde{\Psi}_n^{(0)}\rangle) + \\ E_n^{(2)}\,\langle\Psi_n^{(0)}|\tilde{\Psi}_n^{(0)}\rangle\end{aligned} \tag{B.43}$$

If Ψ and $\tilde{\Psi}$ are biorthogonal over $\hat{H}_0$ then $\langle\Psi_n^{(0)}|\hat{H}_0|\tilde{\Psi}_l^{(0)}\rangle = 0$ and $\langle\Psi_n^{(0)}|\tilde{\Psi}_l^{(0)}\rangle = 0$ if $l \neq n$. This further implies that $\langle\Psi_n^{(l)}|\hat{H}_0|\tilde{\Psi}_n^{(m)}\rangle = 0$ and $\langle\Psi_n^{(l)}|\tilde{\Psi}_n^{(m)}\rangle = 0$ if $l \neq m$, because perturbative corrections to the wavefunction are constructed in a basis of excited state determinants that are biorthogonal to the reference state. This drastically simplifies the above equations:

$$\langle\Psi_n^{(0)}|\hat{H}_0|\tilde{\Psi}_n^{(0)}\rangle = E_n^{(0)}\,\langle\Psi_n^{(0)}|\tilde{\Psi}_n^{(0)}\rangle \tag{B.44}$$

$$\langle\Psi_n^{(0)}|\hat{V}|\tilde{\Psi}_n^{(0)}\rangle = E_n^{(1)}\,\langle\Psi_n^{(0)}|\tilde{\Psi}_n^{(0)}\rangle \tag{B.45}$$

$$\langle\Psi_n^{(0)}|\hat{V}|\tilde{\Psi}_n^{(1)}\rangle + \langle\Psi_n^{(1)}|\hat{V}|\tilde{\Psi}_n^{(0)}\rangle = E_n^{(2)}\,\langle\Psi_n^{(0)}|\tilde{\Psi}_n^{(0)}\rangle \tag{B.46}$$

Finally, it remains to derive first-order corrections to both Ψ_0 and $\tilde{\Psi}_0$, and use those to compute the second-order energy correction. This derivation closely follows that of standard RSPT theory, although now the wavefunction is expanded in a biorthogonal basis, including normalization factor in case biorthogonal basis states are not also normalized:

$$\langle \Psi_n^{(m)}| = \sum_l \frac{\langle \Psi_n^{(m)}|\tilde{\Psi}_l^{(0)}\rangle \langle \Psi_l^{(0)}|}{\langle \Psi_l^{(0)}|\tilde{\Psi}_l^{(0)}\rangle} = \sum_l C_{n,l} \frac{\langle \Psi_l^{(0)}|}{\langle \Psi_l^{(0)}|\tilde{\Psi}_l^{(0)}\rangle} \tag{B.47}$$

$$|\tilde{\Psi}_n^{(m)}\rangle = \sum_l \frac{|\tilde{\Psi}_l^{(0)}\rangle \langle \Psi_l^{(0)}|\tilde{\Psi}_n^{(m)}\rangle}{\langle \Psi_l^{(0)}|\tilde{\Psi}_l^{(0)}\rangle} = \sum_l \frac{|\tilde{\Psi}_l^{(0)}\rangle}{\langle \Psi_l^{(0)}|\tilde{\Psi}_l^{(0)}\rangle} \tilde{C}_{l,n} \tag{B.48}$$

For the first-order RSPT correction to the wavefunction, we need to find biorthogonal expansion coefficients of the form:

$$C_{n,l}^{(1)} = \langle \Psi_n^{(1)}|\tilde{\Psi}_l^{(0)}\rangle \tag{B.49}$$

$$\tilde{C}_{l,n}^{(1)} = \langle \Psi_l^{(0)}|\tilde{\Psi}_n^{(1)}\rangle \tag{B.50}$$

Symmetrically inserting the resolution of the identity for biorthogonal states, $\hat{I} = \sum_l \frac{|\tilde{\Psi}_l^{(0)}\rangle\langle\Psi_l^{(0)}|}{S_l}$ where $S_l = \langle \Psi_l^{(0)}|\tilde{\Psi}_l^{(0)}\rangle$, and using the Einstein summation convention for repeated indices:

$$\begin{aligned} S_l^{-2} \langle \Psi_n^{(0)}|\tilde{\Psi}_l^{(0)}\rangle \langle \Psi_l^{(0)}|\hat{H}_0|\tilde{\Psi}_l^{(0)}\rangle \langle \Psi_l^{(0)}|\tilde{\Psi}_n^{(1)}\rangle + &= E_n^{(0)} S_l^{-1} \langle \Psi_n^{(0)}|\tilde{\Psi}_l^{(0)}\rangle \langle \Psi_l^{(0)}|\tilde{\Psi}_n^{(1)}\rangle + \\ S_l^{-2} \langle \Psi_n^{(1)}|\tilde{\Psi}_l^{(0)}\rangle \langle \Psi_l^{(0)}|\hat{H}_0|\tilde{\Psi}_l^{(0)}\rangle \langle \Psi_l^{(0)}|\tilde{\Psi}_n^{(0)}\rangle + &\quad E_n^{(0)} S_l^{-1} \langle \Psi_n^{(1)}|\tilde{\Psi}_l^{(0)}\rangle \langle \Psi_l^{(0)}|\tilde{\Psi}_n^{(0)}\rangle + \\ S_l^{-1} \frac{1}{2} (\langle \Psi_n^{(0)}|\tilde{\Psi}_l^{(0)}\rangle \langle \Psi_l^{(0)}|\hat{V}|\tilde{\Psi}_n^{(0)}\rangle + \langle \Psi_n^{(0)}|\hat{V}|\tilde{\Psi}_l^{(0)}\rangle \langle \Psi_l^{(0)}|\tilde{\Psi}_n^{(0)}\rangle) &\quad E_n^{(1)} S_l^{-1} \langle \Psi_n^{(0)}|\Psi_l^{(0)}\rangle \langle \Psi_l^{(0)}|\tilde{\Psi}_n^{(0)}\rangle \end{aligned} \tag{B.51}$$

So:

$$\begin{aligned} S_l^{-2} \delta_{n,l} \langle \Psi_l^{(0)}|\hat{H}_0|\tilde{\Psi}_l^{(0)}\rangle \langle \Psi_l^{(0)}|\tilde{\Psi}_n^{(1)}\rangle + &= E_n^{(0)} S_l^{-1} \delta_{n,l} \langle \Psi_l^{(0)}|\tilde{\Psi}_n^{(1)}\rangle + \\ S_l^{-2} \langle \Psi_n^{(1)}|\tilde{\Psi}_l^{(0)}\rangle \langle \Psi_l^{(0)}|\hat{H}_0|\tilde{\Psi}_l^{(0)}\rangle \delta_{n,l} + &\quad E_n^{(0)} S_l^{-1} \langle \Psi_n^{(1)}|\tilde{\Psi}_l^{(0)}\rangle \delta_{l,n} + \\ S_l^{-1} \frac{1}{2} (\delta_{n,l} \langle \Psi_l^{(0)}|\hat{V}|\tilde{\Psi}_n^{(0)}\rangle + \delta_{l,n} \langle \Psi_n^{(0)}|\hat{V}|\tilde{\Psi}_l^{(0)}\rangle) &\quad E_n^{(1)} S_l^{-1} \delta_{n,l}\delta_{l,n} \end{aligned} \tag{B.52}$$

Evaluating the integrals over H_0 according to equation (B.41) yields:

$$\begin{aligned} S_l^{-1} E_l^{(0)} \langle \Psi_l^{(0)} | \tilde{\Psi}_n^{(1)} \rangle + = E_n^{(0)} S_l^{-1} \langle \Psi_l^{(0)} | \tilde{\Psi}_n^{(1)} \rangle \; + \\ S_l^{-1} \langle \Psi_n^{(1)} | \tilde{\Psi}_l^{(0)} \rangle E_l^{(0)} \; + \quad E_n^{(0)} S_l^{-1} \langle \Psi_n^{(1)} | \tilde{\Psi}_l^{(0)} \rangle \; + \\ S_l^{-1} \frac{1}{2} (\langle \Psi_l^{(0)} | \hat{V} | \tilde{\Psi}_n^{(0)} \rangle + \langle \Psi_n^{(0)} | \hat{V} | \tilde{\Psi}_l^{(0)} \rangle) \end{aligned} \tag{B.53}$$

And simplifying:

$$E_l^{(0)}(\langle \Psi_l^{(0)} | \tilde{\Psi}_n^{(1)} \rangle + \langle \Psi_n^{(1)} | \tilde{\Psi}_l^{(0)} \rangle) + \frac{1}{2}(\langle \Psi_l^{(0)} | \hat{V} | \tilde{\Psi}_n^{(0)} \rangle + \langle \Psi_n^{(0)} | \hat{V} | \tilde{\Psi}_l^{(0)} \rangle) = E_n^{(0)}(\langle \Psi_l^{(0)} | \tilde{\Psi}_n^{(1)} \rangle + \langle \Psi_n^{(1)} | \tilde{\Psi}_l^{(0)} \rangle) \tag{B.54}$$

Collecting terms with like indexing:

$$E_l^{(0)} \langle \Psi_l^{(0)} | \tilde{\Psi}_n^{(1)} \rangle + \frac{1}{2} \langle \Psi_l^{(0)} | \hat{V} | \tilde{\Psi}_n^{(0)} \rangle = E_n^{(0)} \langle \Psi_l^{(0)} | \tilde{\Psi}_n^{(1)} \rangle \tag{B.55}$$

$$E_l^{(0)} \langle \Psi_n^{(1)} | \tilde{\Psi}_l^{(0)} \rangle + \frac{1}{2} \langle \Psi_n^{(0)} | \hat{V} | \tilde{\Psi}_l^{(0)} \rangle = E_n^{(0)} \langle \Psi_n^{(1)} | \tilde{\Psi}_l^{(0)} \rangle \tag{B.56}$$

$$\tag{B.57}$$

So, finally:

$$\langle \Psi_l^{(0)} | \tilde{\Psi}_n^{(1)} \rangle = \frac{1}{2} \frac{\langle \Psi_l^{(0)} | \hat{V} | \tilde{\Psi}_n^{(0)} \rangle}{(E_n^{(0)} - E_l^{(0)})} = \tilde{C}_{l,n} \tag{B.58}$$

$$\langle \Psi_n^{(1)} | \tilde{\Psi}_l^{(0)} \rangle = \frac{1}{2} \frac{\langle \Psi_n^{(0)} | \hat{V} | \tilde{\Psi}_l^{(0)} \rangle}{(E_n^{(0)} - E_l^{(0)})} = C_{n,l} \tag{B.59}$$

Combining equations (B.59), (B.58), (B.47), (B.48) and inserting the resultant expressions into (B.46) gives the non-orthogonal second-order perturbation theory energy correction:

$$\sum_l \frac{\langle \Psi_n^{(0)} | \hat{V} | \tilde{\Psi}_l^{(0)} \rangle \langle \Psi_l^0 | \hat{V} | \tilde{\Psi}_n^{(0)} \rangle}{(E_n^{(0)} - E_l^{(0)}) \langle \Psi_l^{(0)} | \tilde{\Psi}_l^{(0)} \rangle} = E_n^{(2)} \langle \Psi_n^{(0)} | \tilde{\Psi}_n^{(0)} \rangle \tag{B.60}$$

If the biorthogonal states are also normalized, all the overlap integral terms evaluate to 1, simplifying this equation to closely resemble the 'standard' RSPT2 expression:

$$\sum_{l} \frac{\langle \Psi_n^{(0)} | \hat{V} | \tilde{\Psi}_l^{(0)} \rangle \langle \Psi_l^0 | \hat{V} | \tilde{\Psi}_n^{(0)} \rangle}{(E_n^{(0)} - E_l^{(0)})} = E_n^{(2)} \tag{B.61}$$

Note that the order of indices within the integral equations is important when $\hat{V}$ is not a Hermitian operator. In such cases:

$$\langle \Psi_n^{(0)} | \hat{V} | \tilde{\Psi}_l^{(0)} \rangle \langle \Psi_l^0 | \hat{V} | \tilde{\Psi}_n^{(0)} \rangle \neq | \langle \Psi_n^{(0)} | \hat{V} | \tilde{\Psi}_l^{(0)} \rangle |^2 \tag{B.62}$$

To make further progress, an explicit form for – and partitioning of – the Hamiltonian must be defined. Two options are possible: we could choose $\hat{H}_0$ to be the one-electron Fock operator and $\hat{V}$ to be the difference between the true and mean-field Coulomb potentials. Alternatively, we could choose $\hat{H}_0$ to be the one-electron core Hamiltonian (for biorthogonalized orbitals) and $\hat{V}$ to be the Coulomb operator. In either case, the CI matrix element energy is equal to the sum of 0th and 1st order MP2 terms.

$$\frac{\langle \Psi_0^{(0)} | \hat{H} | \tilde{\Psi}_0^{(0)} \rangle}{\langle \Psi_0^{(0)} | \tilde{\Psi}_0^{(0)} \rangle} = \frac{\langle \Psi_0^{(0)} | \hat{H}_0 + \hat{V} | \tilde{\Psi}_0^{(0)} \rangle}{\langle \Psi_0^{(0)} | \tilde{\Psi}_0^{(0)} \rangle} = E_0^{(0)} + E_0^{(1)} \tag{B.63}$$

Similarly, both wavefunctions are represented by Hartree-Fock determinants: $\Psi_0^{(0)} = \Phi$ and $\Phi_0^{(0)} = \tilde{\Phi}$, and perturbative wavefunction corrections are expanded in a basis of excited state determinants:

$$\langle \Psi_0^{(1)} | = \sum_{i,a} C_{i,a}^{(1)} \langle \Phi_i^a | + \sum_{i<j,a<b} C_{i,j,a,b}^{(1)} \langle \Phi_{i,j}^{a,b} | + \dots \tag{B.64}$$

$$| \tilde{\Psi}_0^{(1)} \rangle = \sum_{i,a} C_{i,a}^{(1)} | \tilde{\Phi}_i^a \rangle + \sum_{i<j,a<b} C_{i,j,a,b}^{(1)} | \tilde{\Phi}_{i,j}^{a,b} \rangle + \dots \tag{B.65}$$

Now, i, j, a, b are indices for excited states, rather than l. Substituting these expressions into equation (B.61), and representing the reference states also by biorthogonalized HF determinants, $\Psi_n^{(0)} = \Phi$ and $\tilde{\Psi}_n^{(0)} = \tilde{\Phi}$ yields the biorthogonal MP2 energy expression (assuming biorthogonal states are normalized, otherwise normalization factors need to be included as per equation (B.60)):

$$\sum_{i,a} \frac{\langle \Phi | \hat{V} | \tilde{\Phi}_i^a \rangle \langle \Phi_i^a | \hat{V} | \tilde{\Phi} \rangle}{(E_0 - E_i^a)} + \frac{1}{4} \sum_{i,j,a,b} \frac{\langle \Phi | \hat{V} | \tilde{\Phi}_{i,j}^{a,b} \rangle \langle \Phi_{i,j}^{a,b} | \hat{V} | \tilde{\Phi} \rangle}{(E_0 - E_{i,j}^{a,b})} = E_n^{(2)} \tag{B.66}$$

The precise evaluation of the terms in the above equation depends on the partitioning chosen. By far the simplest choice is to set $\hat{H}_0$ to be the core Hamiltonian and $\hat{V}$ to be the Coulomb operator. Then each of the terms in the above equation can be readily evaluated using the two-electron terms from the modified Slater-Condon

rules presented in Chapter 2:

$$\begin{aligned}
\langle\Phi|\hat{V}|\widetilde{\Phi}\rangle_I &= S^{\mathrm{red}} \sum_J \frac{\langle IJ||IJ\rangle}{s_J} \\
\langle\Phi|\hat{V}|\widetilde{\Phi}\rangle_{IJ} &= S^{\mathrm{red}} \langle IJ||IJ\rangle \\
\langle\Phi|\hat{V}|\widetilde{\Phi}\rangle_{IJK} &= 0
\end{aligned} \tag{B.67}$$

where I, J and K index the non-orthogonal orbital-pairs between the biorthogonalised input determinants, and all other terms have the same meanings as defined in Chapter 2.

Of course, the $i \rightarrow a$ and $j \rightarrow b$ excitations specified in equation (B.66) must be made before each determinant is biorthogonalised against the other reference state.

Physically, the one-electron terms in the MP2 energy expression account for changes to the mean-field electron repulsion potential experienced by an electron in one of the reference state molecular orbitals due to non-orthogonality of the other state's virtual orbitals upon excitation of one electron into that virtual orbital manifold. This is expected to make a very small contribution to the overall MP2 energy correction.

The two-electron terms, on the other hand, are the NOCI equivalent of the electron repulsion integrals that appear in the conventional MP2 energy expression. However, it is important to note that these terms now account for the Coulomb weighted overlap between the ground state of one reference determinant and the doubly-excited states of the other reference determinant.

As such, evaluating these terms requires separate bi-orthogonalisation procedures for each doubly-excited state. For those terms that have two non-zero overlaps

within the biorthogonal set, the required integrals must be evaluated using the MO to AO transformation procedure that underpins the evaluation of all two-electron molecular integrals required by all second order perturbation theories and scales as $O(N_{\text{orbs}}^5)$. This transformation cannot be done once and for all on a single set of orthogonal orbitals. This means that it would be very computationally costly to directly evaluating equation (B.66) using this approach.

Bibliography

[1] Pople, J. A. *Reviews of Modern Physics* **1999**, *71*, 1267–1274.

[2] Iczkowski, R. P.; Margrave, J. L. *Journal of the American Chemical Society* **1961**, *83*, 3547–3551.

[3] Grove, N. P.; Cooper, M. M.; Cox, E. L. *Journal of Chemical Education* **2012**, *89*, 850–853.

[4] Löwdin, P.-O. *Physical Review* **1955**, *97*, 1490–1508.

[5] Pople, J. A.; Beveridge, D. L. *Approximate Molecular Orbital Theory*; McGraw-Hill Book Company, 1970; pp 85–159.

[6] Haim, A. *Journal of Chemical Education* **1991**, *68*, 737–738.

[7] Weinhold, F. *Journal of Chemical Education* **1999**, *76*, 1141–1146.

[8] Klopman, G. *Journal of the American Chemical Society* **1968**, *90*, 223–234.

[9] Bradley, J. D.; Gerrans, G. C. *Journal of Chemical Education* **1973**, *50*, 463–466.

[10] Jensen, F. *Introduction to Computational Chemistry*, 2nd ed.; John Wiley & Sons, Ltd, 2013; pp 494–496.

[11] Slater, J. C. *Physical Review* **1929**, *34*, 1293–1322.

[12] Löwdin, P. O. *Physical Review* **1955**, *97*, 1509–1520.

[13] Bartlett, R. J.; Stanton, J. F. *Reviews in Computational Chemistry*; John Wiley & Sons, Ltd, 2007; pp 65–169.

[14] Tew, D.; Klopper, W.; Helgaker, T. *Journal of Computational Chemistry* **2007**, *28*, 1307–1320.

[15] Sinanoğlu, O. *Advances in Chemical Physics*; John Wiley & Sons, Ltd, 2007; pp 237–282.

[16] Hollett, J. W.; Gill, P. M. *The Journal of Chemical Physics* **2011**, *134*, 114111/1–5.

[17] Grimme, S.; Hansen, A. *Angewandte Chemie International Edition* **2015**, *54*, 12308–12313.

[18] Stein, C. J.; Von Burg, V.; Reiher, M. *Journal of Chemical Theory and Computation* **2016**, *12*, 3764–3773.

[19] Handy, N. C.; Cohen, A. J. *Molecular Physics* **2001**, *99*, 403–412.

[20] Ramos-Cordoba, E.; Salvador, P.; Matito, E. *Physical Chemistry Chemical Physics* **2016**, *18*, 24015–24023.

[21] Löwdin, P.-O. *Physical Review* **1955**, *97*, 1474–1489.

[22] Sakurai, J. In *Modern Quantum Mechanics*; Tuan, S. F., Ed.; Addison-Wesley, 1967; pp 285–345.

[23] Fink, R. F. *The Journal of Chemical Physics* **2016**, *145*, 184101/1–12.

[24] Helgaker, T.; Gauss, J.; Jørgensen, P.; Olsen, J. *The Journal of Chemical Physics* **2002**, *106*, 6430–6440.

[25] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications, 1982; pp 231–270.

[26] Nakatsuji, H. *The Journal of Chemical Physics* **2000**, *113*, 2949–2956.

[27] Fales, B. S.; Levine, B. G. *Journal of Chemical Theory and Computation* **2015**, *11*, 4708–4716.

[28] Sherrill, D. C.; Schaefer, H. F. I. *Advances in Quantum Chemistry* **1999**, *34*, 143–269.

[29] Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *The Journal of Chemical Physics* **1987**, *87*, 5968–5975.

[30] Csizmadia, I. G.; Sutcliffe, B. T.; Barnett, M. P. *Canadian Journal of Chemistry* **1964**, *42*, 1645–1663.

[31] Gilbert, A. T. B.; Besley, N. A.; Gill, P. M. W. *The Journal of Physical Chemistry. A* **2008**, *112*, 13164–13171.

[32] Rosen, M. *Physical Review A* **1970**, *1*, 1285–1288.

[33] Schmidt, M. W.; Hull, E. A.; Windus, T. L. *Journal of Physical Chemistry A* **2015**, *119*, 10408–10427.

[34] Pople, J. A.; Binkley, J. S.; Seeger, R. *International Journal of Quantum Chemistry* **1976**, *10*, 1–19.

[35] Bartlett, R. J.; Purvis, G. D. *International Journal of Quantum Chemistry* **1978**, *14*, 561–581.

[36] Jiménez-Hoyos, C. A.; Henderson, T. M.; Scuseria, G. E. *Journal of Chemical Theory and Computation* **2011**, *7*, 2667–2674.

[37] Duch, W.; Diercksen, G. H. *The Journal of Chemical Physics* **1994**, *101*, 3018–3030.

[38] Paldus, J.; Li, X. *Advances in Chemical Physics*; John Wiley & Sons, Ltd, 2007; pp 1–175.

[39] Bartlett, R. J. *Molecular Physics* **2010**, *108*, 2905–2920.

[40] Bartlett, R. J. *Annual Review of Physical Chemistry* **1981**, *32*, 359–401.

[41] Cole, S. J.; Bartlett, R. J. *The Journal of Chemical Physics* **2002**, *86*, 873–881.

[42] Barlett, R. J.; Sekino, H.; Purvis, G. D. *Chemical Physics Letters* **1983**, *98*, 66–71.

[43] Čížek, J. *The Journal of Chemical Physics* **1966**, *45*, 4256–4266.

[44] Kvaal, S. *Molecular Physics* **2013**, *111*, 1100–1108.

[45] Krylov, A. I. *Q-Chem User's Manual*; 2018; pp 232–234.

[46] Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chemical Physics Letters* **1989**, *157*, 479–483.

[47] Řezáč, J.; Hobza, P. *Journal of Chemical Theory and Computation* **2013**, *9*, 2151–2155.

[48] Sengupta, A.; Ramabhadran, R. O.; Raghavachari, K. *Journal of Computational Chemistry* **2016**, *37*, 286–295.

[49] Hinze, J. *The Journal of Chemical Physics* **1973**, *59*, 6424–6432.

[50] Schmidt, M. W.; Gordon, M. S. *Annual Review of Physical Chemistry* **1998**, *49*, 233–266.

[51] Veryazov, V.; Malmqvist, P.-Å.; Roos, B. O. *International Journal of Quantum Chemistry* **2011**, *111*, 3329–3338.

[52] Smith, J. E.; Mussard, B.; Holmes, A. A.; Sharma, S. *Journal of Chemical Theory and Computation* **2017**, *13*, 5468–5478.

[53] Bao, J. J.; Dong, S. S.; Gagliardi, L.; Truhlar, D. G. *Journal of Chemical Theory and Computation* **2018**, *14*, 2017–2025.

[54] Darvesh, K. V.; Grein, F. *International Journal of Quantum Chemistry* **1985**, *28*, 247–256.

[55] Rogers, D. M.; Wells, C.; Joseph, M.; Boddington, V. J.; McDouall, J. J. W. *Journal of Molecular Structure: THEOCHEM* **1998**, *434*, 239–245.

[56] Olsen, J. *International Journal of Quantum Chemistry* **2011**, *111*, 3267–3272.

[57] Roos, B. O. In *Advances in Chemical Physics: Ab Initio Methods in Quantum Chemistry Part 2*; Lawley, K. P., Ed.; Advances in Chemical Physics; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1987; Vol. 69; pp 399–445.

[58] Knowles, P. J.; Sexton, G. J.; Handy, N. C. *Chemical Physics* **1982**, *72*, 337–347.

[59] Pulay, P. *International Journal of Quantum Chemistry* **2011**, *111*, 3273–3279.

[60] Krylov, A. I. *Chemical Physics Letters* **2001**, *350*, 522–530.

[61] Rinkevicius, Z.; Vahtras, O.; Ågren, H. *The Journal of Chemical Physics* **2010**, *133*, 114104/1–12.

[62] Casanova, D.; Head-Gordon, M. *The Journal of Chemical Physics* **2008**, *129*, 064104/1–12.

[63] Krylov, A. I. *Accounts of Chemical Research* **2006**, *39*, 83–91.

[64] Thom, A. J. W.; Head-Gordon, M. *The Journal of Chemical Physics* **2009**, *131*, 124113/1–5.

[65] Thom, A. J. W.; Head-Gordon, M. *Physical Review Letters* **2008**, *101*, 193001/1–4.

[66] Krylov, A. I.; Sherrill, D. C.; Byrd, E. F. C.; Head-Gordon, M. *The Journal of Chemical Physics* **1998**, *109*, 10669–10678.

[67] Krylov, A. I. *Q-Chem User's Manual*; 2015; pp 248–250.

[68] Hiberty, P. C.; Shaik, S. *Journal of Computational Chemistry* **2007**, *28*, 137–151.

[69] Goddard, W. A.; Dunning, T. H.; Hunt, W. J.; Hay, P. J. *Accounts of Chemical Research* **1973**, *6*, 368–376.

[70] Lawler, K. V.; Beran, G. J. O.; Head-Gordon, M. *The Journal of Chemical Physics* **2008**, *128*, 24107/1–13.

[71] Langlois, J. M.; Yamasaki, T.; Muller, R. P.; Goddard, W. A. *Journal of Physical Chemistry* **1994**, *98*, 13498–13505.

[72] Voorhis, T. V.; Head-Gordon, M. *Chemical Physics Letters* **2000**, *317*, 575–580.

[73] Wang, Q.; Zou, J.; Xu, E.; Pulay, P.; Li, S. *Journal of Chemical Theory and Computation* **2019**, *15*, 141–153.

[74] Crittenden, D. L. *Journal of Physical Chemistry A* **2013**, *117*, 3852–3860.

[75] Mayhall, N. J.; Horn, P. R.; Sundstrom, E. J.; Head-Gordon, M. *Physical Chemistry Chemical Physics* **2014**, *16*, 22694–22705.

[76] Sundstrom, E. J.; Head-Gordon, M. *The Journal of Chemical Physics* **2014**, *140*, 114103/1–11.

[77] Rogers, D. M.; Wells, C.; Joseph, M.; Boddington, V. J.; McDouall, J. J. W. *Journal of Molecular Structure: THEOCHEM* **1998**, *434*, 239–245.

[78] Casanova, D.; Head-Gordon, M. *Physical Chemistry Chemical Physics* **2009**, *11*, 9779–9790.

[79] Bell, F.; Zimmerman, P. M.; Casanova, D.; Goldey, M.; Head-Gordon, M. *Physical Chemistry Chemical Physics* **2013**, *15*, 358–366.

[80] Tsuchimochi, T.; Scuseria, G. E. *The Journal of Chemical Physics* **2010**, *133*, 141102/1–4.

[81] Tsuchimochi, T.; Henderson, T. M.; Scuseria, G. E.; Savin, A. *The Journal of Chemical Physics* **2010**, *133*, 134108/1–11.

[82] Hurley, A. C.; Lennard-Jones, J.; Pople, J. A. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **1953**, *220*, 446–455.

[83] Beran, G. J. O.; Austin, B.; Sodt, A.; Head-Gordon, M. *The Journal of Physical Chemistry A* **2005**, *109*, 9183–9192.

[84] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: An Introduction to Advanced Electronic Structure Theory*; Dover Publications, 1982; pp 111–145.

[85] Phillips, J. C. *Physical Review* **1961**, *123*, 420–424.

[86] Boys, S. F.; Egerton, A. C. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1950**, *200*, 542–554.

[87] Schlegel, H. B.; Frisch, M. J. *International Journal of Quantum Chemistry* **1995**, *54*, 83–87.

[88] Pulay, P. *Journal of Computational Chemistry* **1982**, *3*, 556–560.

[89] Hamilton, T. P.; Pulay, P. *The Journal of Chemical Physics* **1986**, *84*, 5728–5734.

[90] Amos, T.; Snyder, L. C. *The Journal of Chemical Physics* **1964**, 1773–1783.

[91] Jensen, F. *Introduction to Computational Chemistry*, 2nd ed.; 2007; Chapter 4, pp 148–150.

[92] Roothaan, C. C. J. *Reviews of Modern Physics* **1960**, *32*, 179–185.

[93] Plakhutin, B. N.; Davidson, E. R. *The Journal of Chemical Physics* **2014**, *140*, 014102/1–15.

[94] Plakhutin, B. N.; Gorelik, E. V.; Breslavskaya, N. N. *The Journal of Chemical Physics* **2006**, *125*, 204110/1–10.

[95] Sears, J. S.; Sherrill, C. D.; Krylov, A. I. *The Journal of Chemical Physics* **2003**, *118*, 9084–9094.

[96] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Methods*; Dover Publications, 1982; pp 70–72.

[97] Casanova, D. *The Journal of Chemical Physics* **2012**, *137*, 084105/1–10.

[98] Burrows, B. L.; Cohen, M. *Journal of Physics B: Atomic, Molecular and Optical Physics* **2019**, *52*, 095003/1–7.

[99] Moskowitz, J. W.; Schmidt, K. E.; Lee, M. A.; Kalos, M. H. *The Journal of Chemical Physics* **1982**, *77*, 349–355.

[100] Adams, W. H. *Physical Review* **1962**, *127*, 1650–1658.

[101] Stwalley, W. C.; Zemke, W. T. *Journal of Physical and Chemical Reference Data* **1993**, *22*, 87–112.

[102] Docken, K. K.; Hinze, J. *The Journal of Chemical Physics* **1972**, *57*, 4928–4936.

[103] Wetthasinghe, S. T.; Rassolov, V. A. *Chemical Physics Letters* **2019**, *730*, 112–116.

[104] Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657–2664.

[105] Hehre, W. J.; Ditchfield, R.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1970**, *52*, 2769–2773.

[106] Pritchard, B. P.; Altarawy, D.; Windus, T. L. **2019**, *59*, 4814–4820.

[107] Dunning, T. H. *The Journal of Chemical Physics* **1989**, *90*, 1007–1023.

[108] Prascher, B. P.; Woon, D. E.; Peterson, K. A.; Dunning, T. H.; Wilson, A. K. *Theoretical Chemistry Accounts* **2011**, *128*, 69–82.

[109] Shao, Y. et al. *Molecular Physics* **2015**, *113*, 184–215.

[110] Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; S, K.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.;

Dupuis, M.; Montgomery, J. A. *Journal of Computaional Chemistry* **1993**, *14*, 1347–1363.

[111] Kosman, W. M.; Hinze, J. *Journal of Molecular Spectroscopy* **1975**, *56*, 93–103.

[112] Lopez, X.; Piris, M.; Matxain, J. M.; Ruipérez, F.; Ugalde, J. M. *ChemPhysChem* **2011**, *12*, 1673–1676.

[113] Salem, L.; Rowland, C. *Angewandte Chemie - International Edition* **1972**, *11*, 92–111.

[114] Wallace, R. *Chemical Physics Letters* **1989**, *159*, 35–36.

[115] Sebastian, R.; Wallace, R. *Chemical Physics Letters* **1991**, *177*, 404–406.

[116] Wallace, R. *Chemical Physics* **1990**, *141*, 241–247.

[117] Merer, A. J.; Mulliken, R. S. *Chemical Reviews* **1969**, *69*, 639–656.

[118] Dreuw, A.; Wormit, M. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2015**, *5*, 82–95.

[119] Chen, W. K.; Liu, X. Y.; Fang, W. H.; Dral, P. O.; Cui, G. *Journal of Physical Chemistry Letters* **2018**, *9*, 6702–6708.

[120] Matsika, S.; Krylov, A. I. *Chemical Reviews* **2018**, *118*, 6925–6926.

[121] Chien, A. D.; Zimmerman, P. M. *The Journal of Chemical Physics* **2017**, *146*, 014103/1–12.

[122] Subotnik, J. E. *The Journal of Chemical Physics* **2011**, *135*, 71104/1–4.

[123] Mochizuki, Y.; Koikegami, S.; Amari, S.; Segawa, K.; Kitaura, K.; Nakano, T. *Chemical Physics Letters* **2005**, *406*, 283–288.

[124] Molina, V.; Smith, B. R.; Merchán, M. *Chemical Physics Letters* **1999**, *309*, 486–494.

[125] Head-Gordon, M.; Rico, R. J.; Oumi, M.; Lee, T. J. *Chemical Physics Letters* **1994**, *219*, 21–29.

[126] Goerigk, L.; Grimme, S. *The Journal of Chemical Physics* **2010**, *132*, 184103/1–9.

[127] Cave, R. J. In *Modern Electronic Structure Theory and Applications in Organic Chemistry*; Davidson, E. R., Ed.; WORLD SCIENTIFIC, 1997; pp 197–255.

[128] Hirata, S.; Head-Gordon, M. *Chemical Physics Letters* **1999**, *302*, 375–382.

[129] Sneskov, K.; Christiansen, O. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 566–584.

[130] Christiansen, O.; Jørgensen, P.; Hättig, C. *International Journal of Quantum Chemistry* **1998**, *68*, 1–52.

[131] Krylov, A. I. *Annual Review of Physical Chemistry* **2008**, *59*, 433–462.

[132] González, L.; Escudero, D.; Serrano-Andrés, L. *ChemPhysChem* **2012**, *13*, 28–51.

[133] Bulik, I. W.; Henderson, T. M.; Scuseria, G. E. *Journal of Chemical Theory and Computation* **2015**, *11*, 3171–3179.

[134] Roos, B. O. *Accounts of Chemical Research* **1999**, *32*, 137–144.

[135] Shu, Y.; Hohenstein, E. G.; Levine, B. G. *The Journal of Chemical Physics* **2015**, *142*, 024102/1–14.

[136] Almlöf, J.; Nemukhin, A. V.; Heiberg, A. *International Journal of Quantum Chemistry* **1981**, *20*, 655–667.

[137] Nangia, S.; Truhlar, D. G.; McGuire, M. J.; Piecuch, P. *The Journal of Physical Chemistry A* **2005**, *109*, 11643–11646.

[138] Grein, F.; Banerjee, A. *International Journal of Quantum Chemistry* **1975**, *9*, 147–154.

[139] Meyer, B.; Domingo, A.; Krah, T.; Robert, V. *Dalton Trans.* **2014**, *43*, 11209–11215.

[140] Liu, X.; Subotnik, J. E. *Journal of Chemical Theory and Computation* **2014**, *10*, 1004–1020.

[141] Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Journal of Physical Chemistry* **1992**, *96*, 135–149.

[142] Oumi, M.; Maurice, D.; Lee, T. J.; Head-Gordon, M. *Chemical Physics Letters* **1997**, *279*, 151–157.

[143] Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. *The Journal of Chemical Physics* **2004**, *121*, 9793–9802.

[144] Hirata, S. *The Journal of Chemical Physics* **2005**, *122*, 94105/1–10.

[145] Vidal, C. R.; Stwalley, W. C. *The Journal of Chemical Physics* **1982**, *77*, 883–898.

[146] Chan, Y. C.; Harding, D. R.; Stwalley, W. C.; Vidal, C. R. *The Journal of Chemical Physics* **1986**, *85*, 2436–2444.

[147] Bennett, O. J.; Dickinson, A. S.; Leininger, T.; Gadéa, F. X. *Monthly Notices of the Royal Astronomical Society* **2003**, *341*, 361–368.

[148] Coxon, J. A.; Dickinson, C. S. *The Journal of Chemical Physics* **2004**, *121*, 9378–9388.

[149] Mok, D. K. W.; Neumann, R.; Handy, N. C. *Journal of Physical Chemistry* **1996**, *100*, 6225–6230.

[150] Ben-Shlomo, S.; Kaldor, U. *The Journal of Chemical Physics* **1988**, *89*, 956–958.

[151] Gadéa, F. X.; Leininger, T. *Theoretical Chemistry Accounts* **2006**, *116*, 566–575.

[152] Holka, F.; Szalay, P. G.; Fremont, J.; Rey, M.; Peterson, K. A.; Tyuterev, V. G. *The Journal of Chemical Physics* **2011**, *134*, 94306/1–14.

[153] Mulliken, R. S. *Physical Review* **1936**, *50*, 1028–1040.

[154] Boutalib, A.; Gadéa, F. X. *The Journal of Chemical Physics* **1992**, *97*, 1144–1156.

[155] Bande, A.; Nakashima, H.; Nakatsuji, H. *Chemical Physics Letters* **2010**, *496*, 347–350.

[156] Barbatti, M.; Paier, J.; Lischka, H. *The Journal of Chemical Physics* **2004**, *121*, 11614–11624.

[157] Feller, D.; Peterson, K. A.; Davidson, E. R. *The Journal of Chemical Physics* **2014**, *141*, 104302/1–21.

[158] Dunning, T. H.; Hunt, W. J.; Goddard, W. A. *Chemical Physics Letters* **1969**, *4*, 147–150.

[159] Ben-Nun, M.; Martinez, T. J. *Chemical Physics Letters* **1998**, *298*, 57–65.

[160] Mulliken, R. S. *The Journal of Chemical Physics* **1977**, *66*, 2448–2451.

[161] Wiberg, K. B.; Wang, Y. G. *Journal of Physical Chemistry A* **2004**, *108*, 9417–9422.

[162] Buenker, R. J.; Bonačić-Koutecký, V.; Pogliani, L. *The Journal of Chemical Physics* **1980**, *73*, 1836–1849.

[163] Ryan, J. A.; Whitten, J. L. *Chemical Physics Letters* **1972**, *15*, 119–123.

[164] Yamamoto, S.; Tatewaki, H. *Chemical Physics* **2003**, *295*, 47–62.

[165] Williams, B. A.; Cool, T. A. *The Journal of Chemical Physics* **1991**, *94*, 6358–6366.

[166] Benavides-Riveros, C. L.; Lathiotakis, N. N.; Marques, M. A. L. *Physical Chemistry Chemical Physics* **2017**, *19*, 12655–12664.

[167] Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *The Journal of Chemical Physics* **1992**, *96*, 1218–1226.

[168] Møller, C.; Plesset, M. S. *Physical Review* **1934**, *46*, 618–622.

[169] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: An Introduction to Advanced Electronic Structure Theory*; Dover Publications, 1982; pp 350–352.

[170] Chaudhuri, R. K.; Freed, K. F.; Hose, G.; Piecuch, P.; Kowalski, K.; Włoch, M.; Chattopadhyay, S.; Mukherjee, D.; Rolik, Z.; Szabados, Á.; Tóth, G.; Surján, P. R. *The Journal of Chemical Physics* **2005**, *122*, 134105/1–9.

[171] Meissner, L. *Molecular Physics* **2006**, *104*, 2073–2083.

[172] Karadakov, P. *International Journal of Quantum Chemistry* **1985**, *27*, 699–707.

[173] Yost, S. R.; Kowalczyk, T.; Van Voorhis, T. *The Journal of Chemical Physics* **2013**, *139*, 174104/1–9.

[174] McDouall, J. J. W.; Peasley, K.; Robb, M. A. *Chemical Physics Letters* **1988**, *148*, 183–189.

[175] Grimme, S.; Waletzke, M. *Physical Chemistry Chemical Physics* **2000**, *2*, 2075–2081.

[176] Nakano, H. *The Journal of Chemical Physics* **1993**, *99*, 7983–7992.

[177] Stanton, J.; Gauss, J.; Cheng, L.; Harding, M.; Matthews, D.; Szalay, P.; Auer, A.; Bartlett, R.; Benedikt, U.; Berger, C. CFOUR, Coupled-Cluster

techniques for Computational Chemistry, a quantum-chemical program package. 2010; http://slater.chemie.uni-mainz.de/cfour/.

[178] Douglas, J. E.; Rabinovitch, B. S.; Looney, F. S. *The Journal of Chemical Physics* **1955**, *23*, 315–323.

[179] Zen, A.; Coccia, E.; Luo, Y.; Sorella, S.; Guidoni, L. *Journal of Chemical Theory and Computation* **2014**, *10*, 1048–1061.

[180] Sension, R. J.; Mayne, L.; Hudson, B. *Journal of the American Chemical Society* **1987**, *109*, 5036–5038.

[181] Bardo, R. D.; Ruedenberg, K. *The Journal of Chemical Physics* **1974**, *60*, 918–931.

[182] Watts, J. D.; Gwaltney, S. R.; Bartlett, R. J. *The Journal of Chemical Physics* **1996**, *105*, 6979–6988.

[183] Yost, S. R.; Head-Gordon, M. *The Journal of Chemical Physics* **2016**, *145*, 54105/1–11.

[184] Oliphant, T. E. *A Guide to Numpy*; Trelgol Publishing USA, 2006; pp 1–378.

[185] Virtanen, P. et al. *Nature Methods* **2020**, *17*, 261–272.

[186] Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chemical Physics Letters* **1988**, *153*, 503–506.

[187] Baker, J.; Wolinski, K. *Journal of Computational Chemistry* **2011**, *32*, 3304–3312.

[188] Kozlowski, P. M.; Davidson, E. R. *The Journal of Chemical Physics* **1994**, *100*, 3672–3682.

[189] Finley, J. P.; Hirao, K. *Chemical Physics Letters* **2000**, *328*, 60–66.

[190] Nakano, H.; Nakatani, J.; Hirao, K. *The Journal of Chemical Physics* **2001**, *114*, 1133–1141.

[191] Roos, B. O.; Andersson, K.; Fülscher, M. P.; Malmqvist, P.-Å.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. *Advances in Chemical Physics: New Methods in Computational Quantum Mechanics, Volume 93* **1996**, *93*, 219–331.

[192] Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. *The Journal of Chemical Physics* **2001**, *114*, 10252–10264.

[193] Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. *Chemical Physics Letters* **2001**, *350*, 297–305.